Analysis Methods for Hadron Colliders III

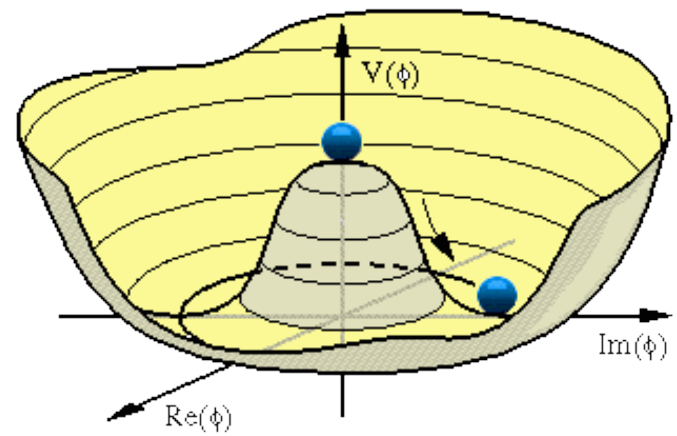
Beate Heinemann

UC Berkeley and Lawrence Berkeley National Laboratory

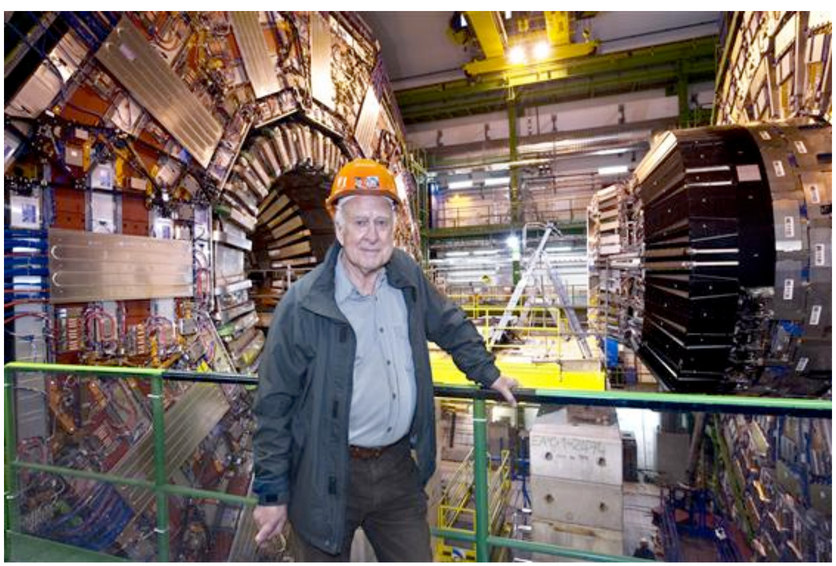
TRIUMF Summer Institute, Vancouver, July 2009

Outline

- Lecture I:
  - Measuring a cross section
    - focus on acceptance
- Lecture II:
  - Searching for a new particle
    - focus on backgrounds
- Lecture III:
  - Continuation on Lecture II (Higgs boson search)
  - Measuring a property of a known particle

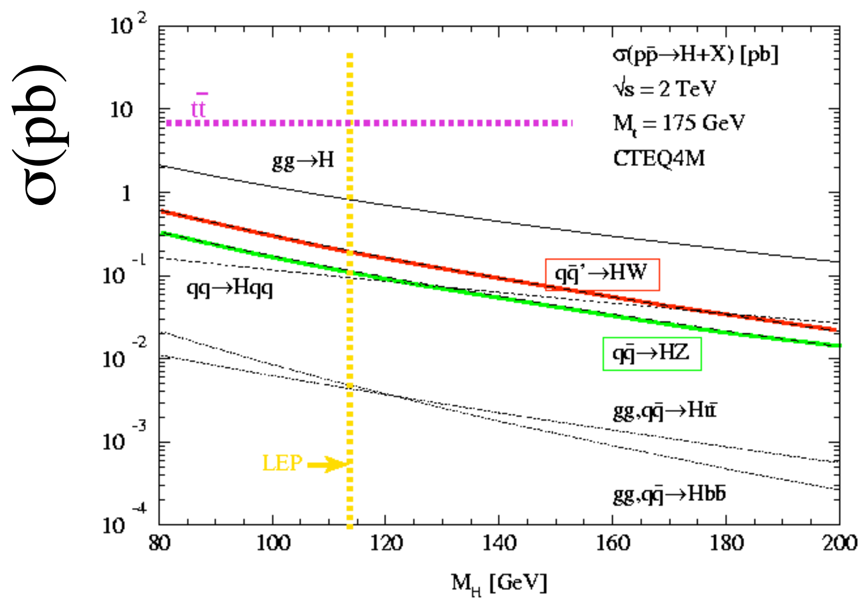


The Higgs Boson

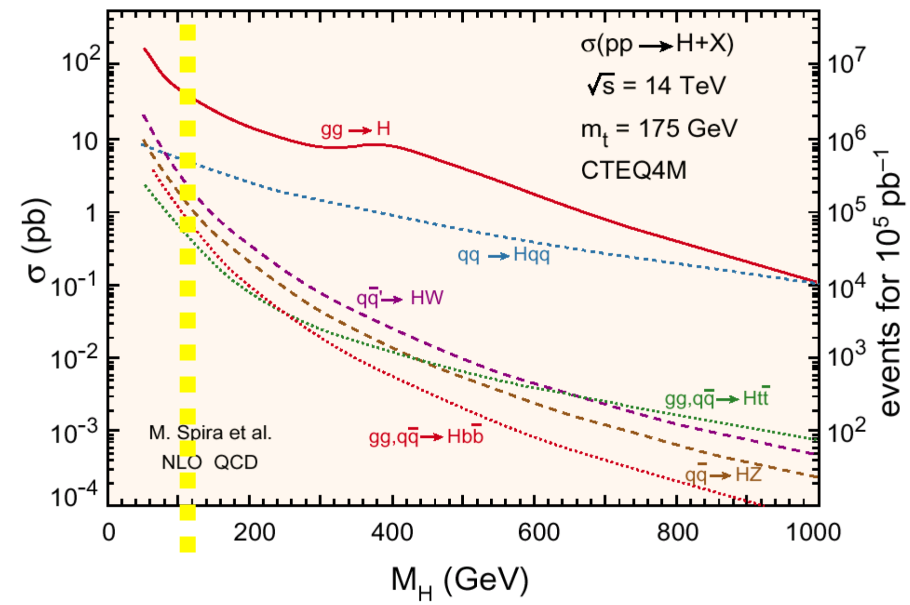


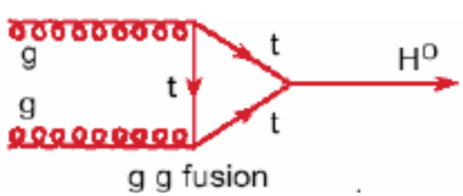
Higgs Production: Tevatron and LHC

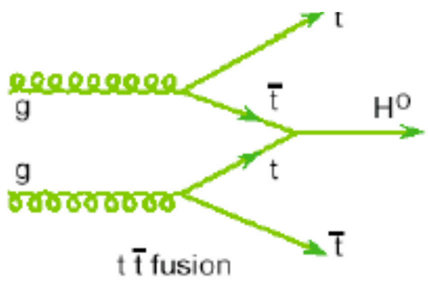
Tevatron

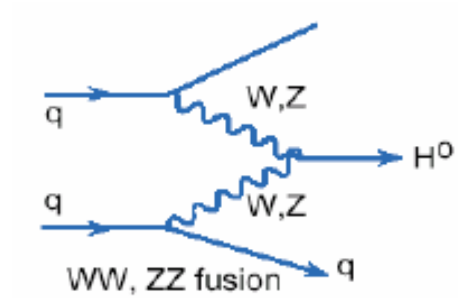


LHC





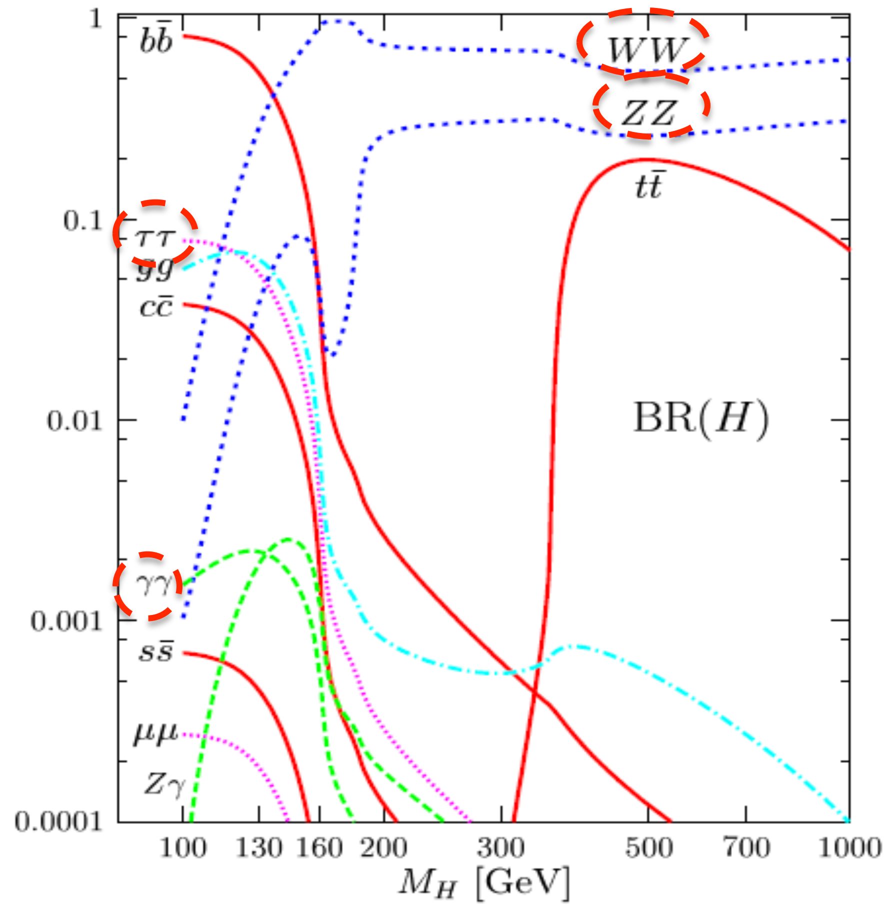




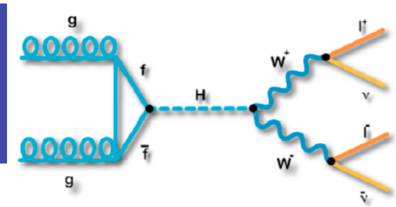
dominant: gg→ H, subdominant: HW, HZ, Hqq

Higgs Boson Decay

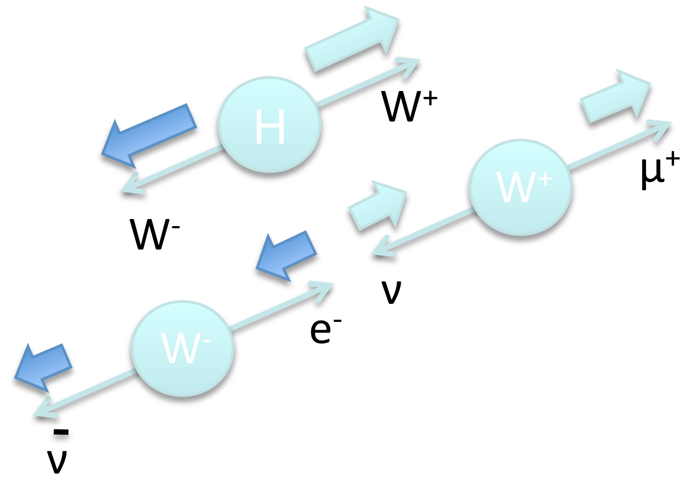
- Depends on Mass
- M_H<130 GeV/c^2:
  - $b\bar{b}$ dominant
  - WW and $\tau\tau$ subdominant
  - $\gamma\gamma$ small but useful
- M_H>130 GeV/c^2:
  - WW dominant
  - ZZ cleanest



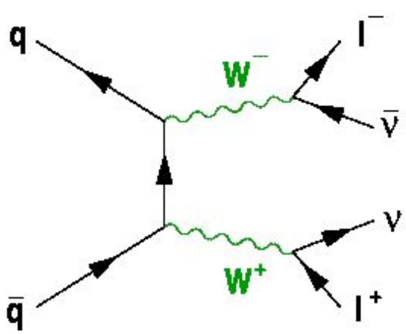
$H \rightarrow WW^{(*)} \rightarrow l^+l^-\bar{\nu}\nu$

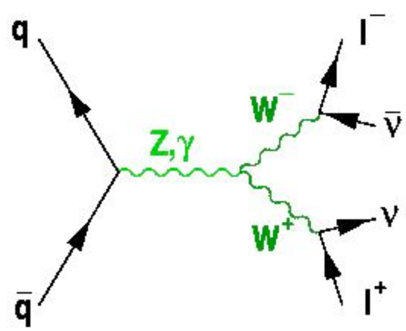


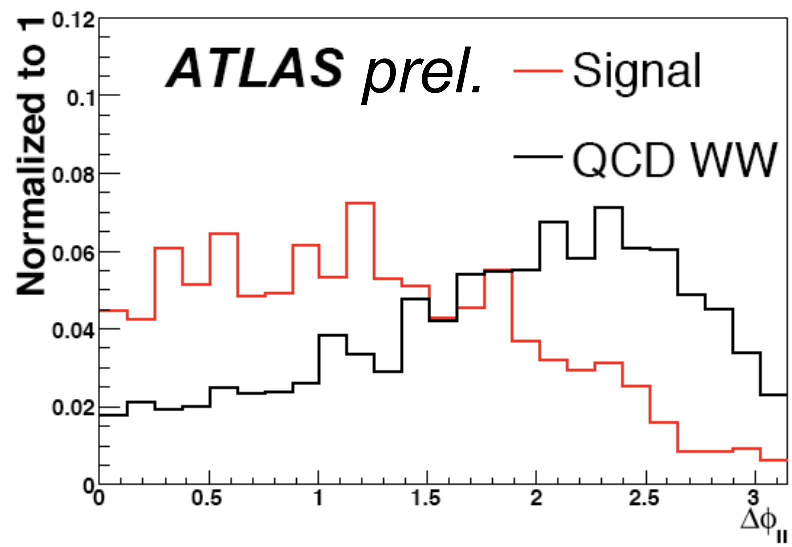
- Higgs mass reconstruction impossible due to two neutrinos in final state
- Make use of spin correlations to suppress WW background:
  - Higgs is scalar: spin=0
  - leptons in $H \rightarrow WW^{(*)} \rightarrow l^+l^-\nu\nu$ are collinear



- Main background:
  - WW production

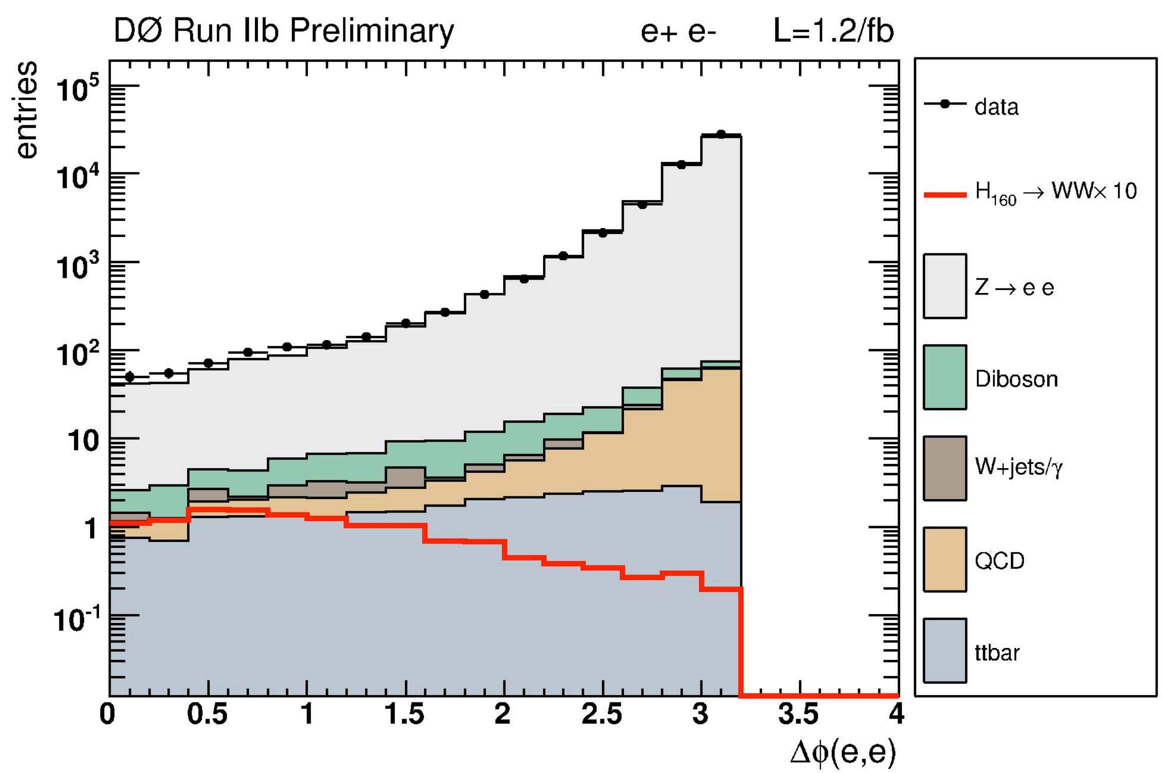






$H \rightarrow WW^{(*)} \rightarrow l^+l^-\nu\nu$ ($l=e,\mu$)

- **Event selection:**
  - 2 isolated e/μ :
    - p_T > 15, 10 GeV
  - Missing E_T >20 GeV
  - Veto on
    - Z resonance
    - Energetic jets



- **Main backgrounds**
  - SM WW production
  - Top
  - Drell-Yan
  - Fake leptons

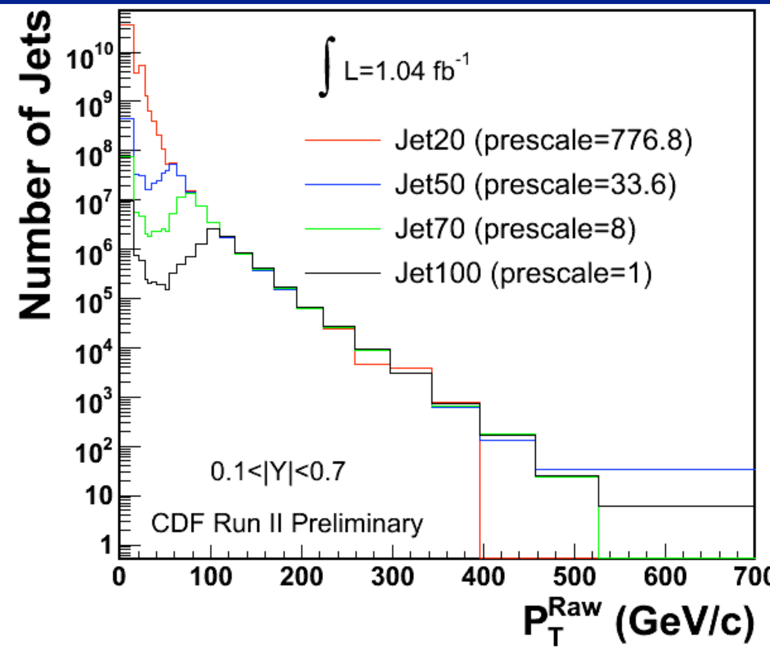
- **Plot everything under the sun**
  - to convince yourself you have the background right

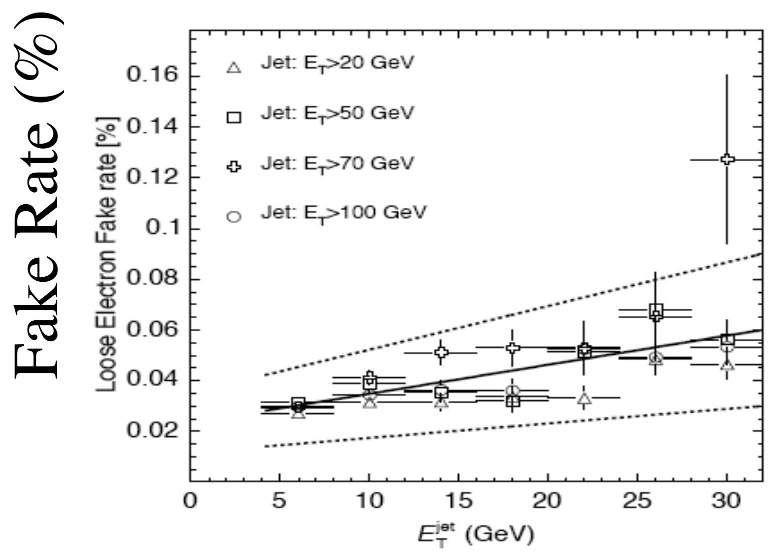
Jets faking Electrons

- Jets can pass electron ID cuts,
  - Mostly due to
    - early showering charged pions
    - Conversions: $\pi^0 \rightarrow \gamma\gamma \rightarrow ee+X$
    - Semileptonic b-decays
  - Difficult to model in MC
    - Hard fragmentation
    - Detailed simulation of calorimeter and tracking volume
- Measured in inclusive jet data at various E_T thresholds
  - Prompt electron content negligible:
    - N_{jet}~10 billion at 50 GeV!
  - **Fake rate per jet:**

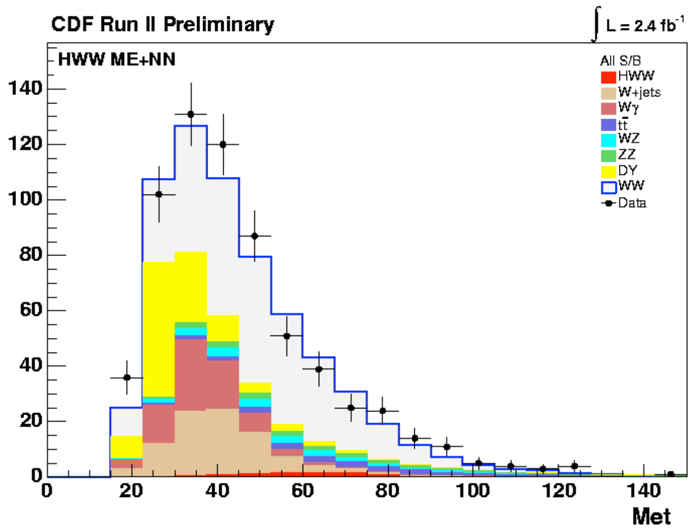
	CDF	ATLAS
Loose cuts	$5x10^{-4}$	$5x10^{-3}$
Tight cuts	$1x10^{-4}$	$1x10^{-5}$

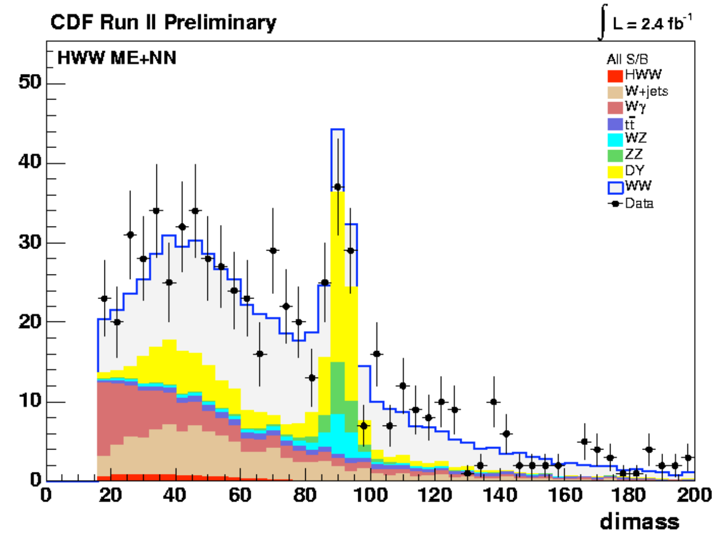
  - Typical uncertainties 50%

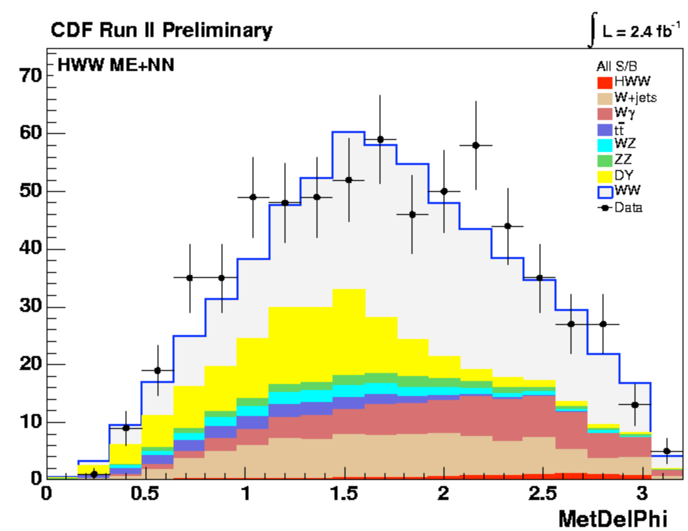


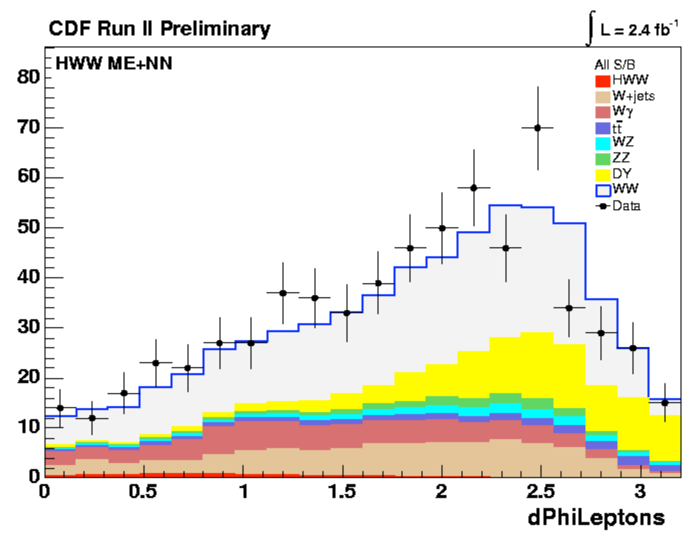


Plot Everything Under the Sun..





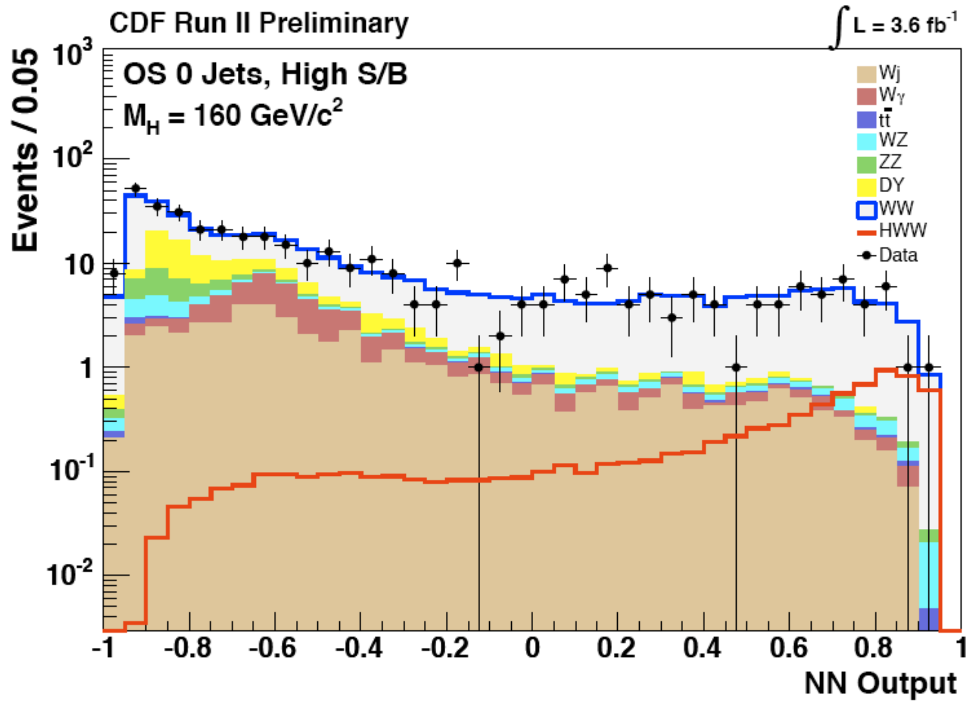




- **Validates the background prediction**
  - **Very often these plots "don't work" since there is some problem**
  - **Now plug all into sophisticated techniques!**

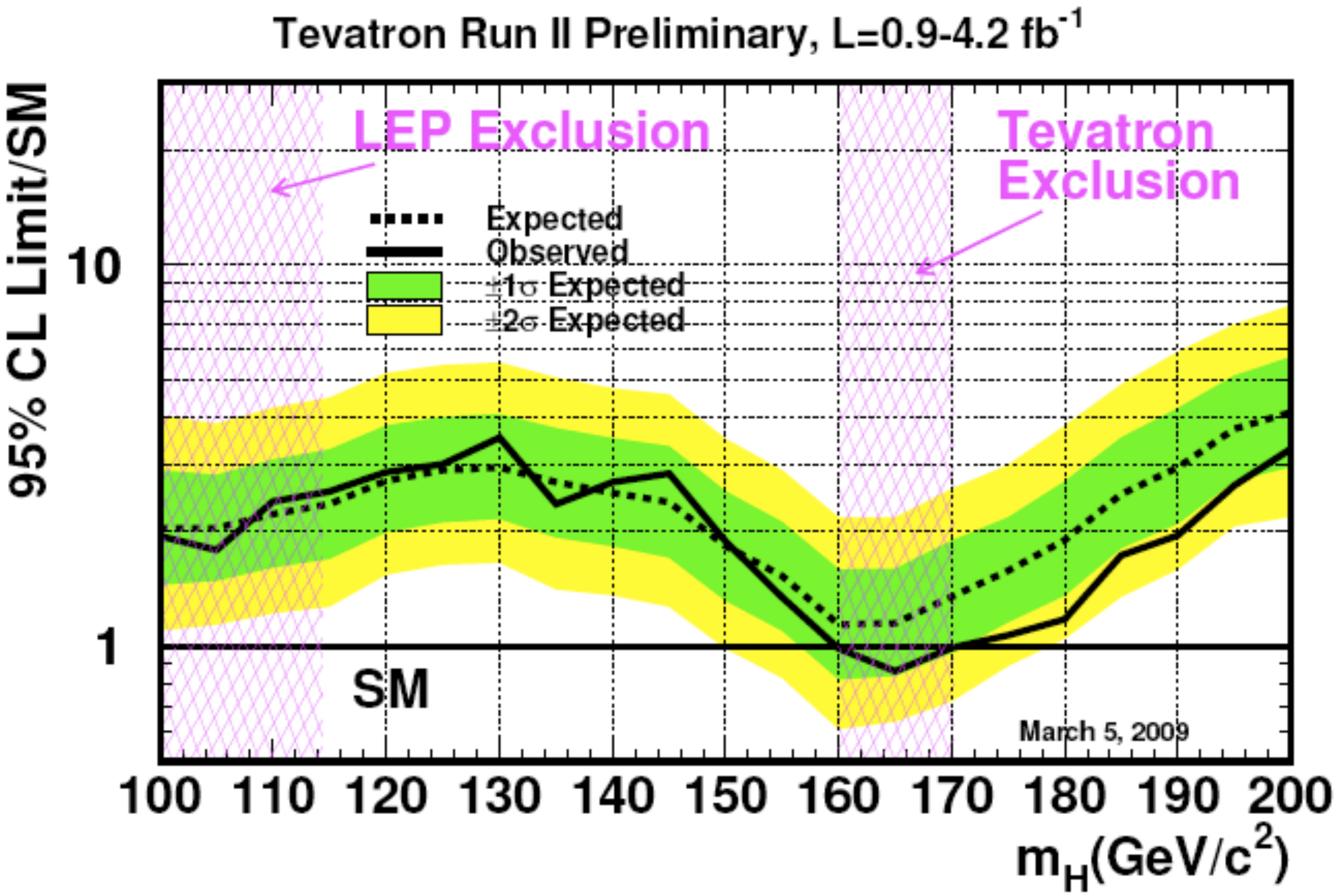
NN Output

$M_H = 160\ \mathrm{GeV}/c^2$	
$t\bar{t}$	1.35 ± 0.21
DY	80 ± 18
WW	318 ± 35
WZ	14 ± 1.9
ZZ	20.7 ± 2.8
W+jets	113 ± 27
$W\gamma$	92 ± 25
Total Background	$\mathbf{637 \pm 67}$
$gg \rightarrow H$	9.5 ± 1.4
Total Signal	$\mathbf{9.5 \pm 1.4}$
Data	**654**



- Data agree well with background hypothesis
  - S/B ~0.3 at high NN values

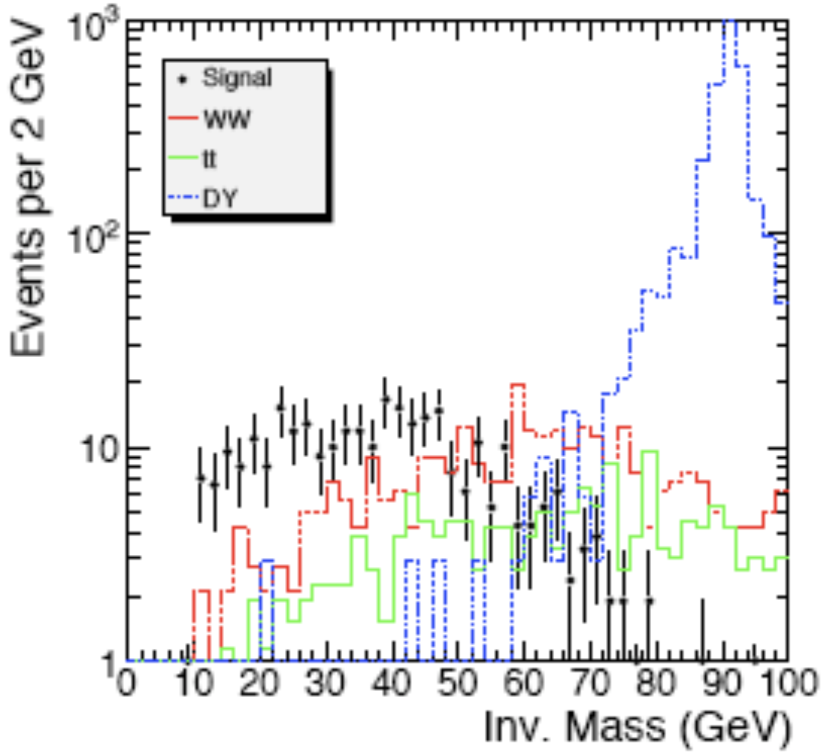
Higgs Cross Section Limit



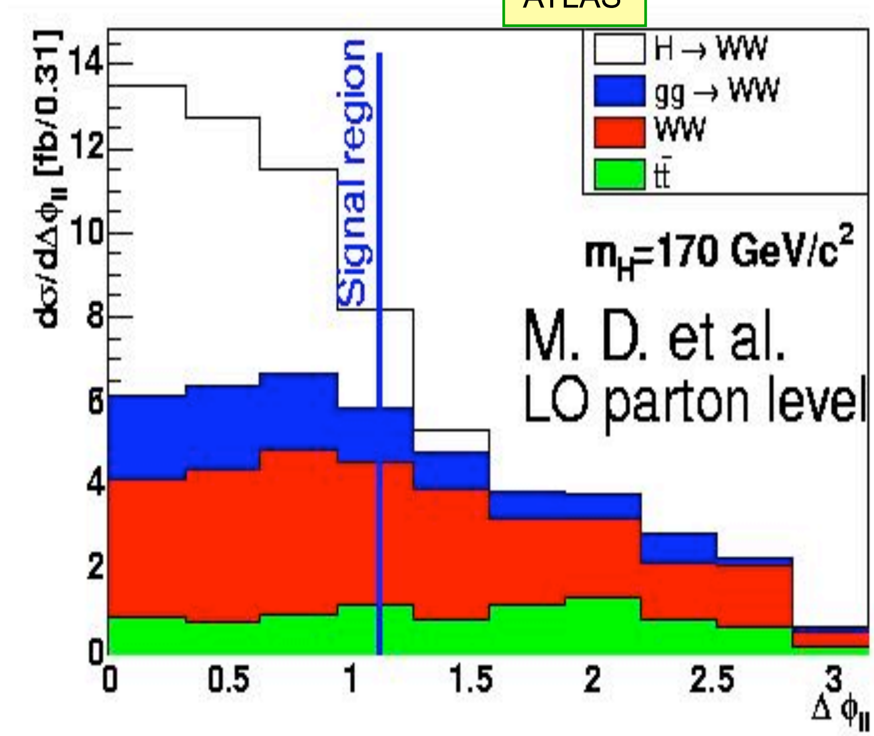
- 160 < m_H < 170 GeV excluded at 95% C.L.
  - Note that the limit is ~1σ better than expected
- For m_H=120 GeV: $\sigma_{limit} / \sigma_{SM}$ = 2.8

Early Higgs Signals at LHC

H→WW* (m_H=170 GeV)

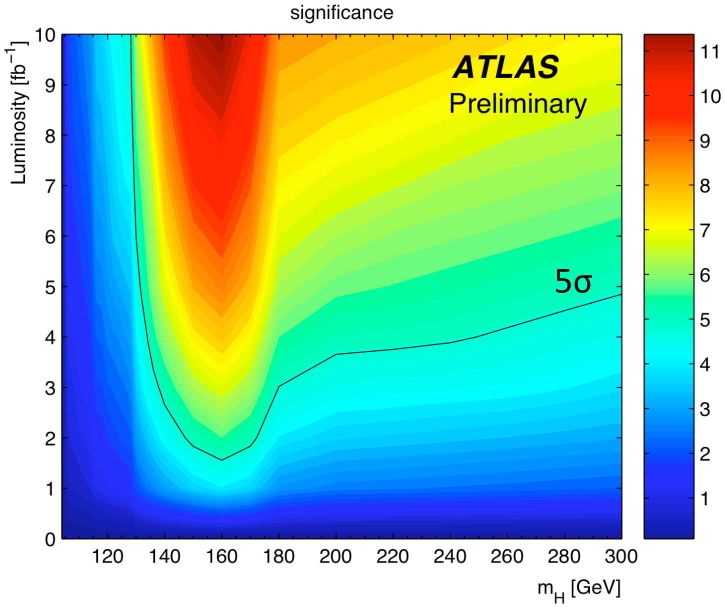


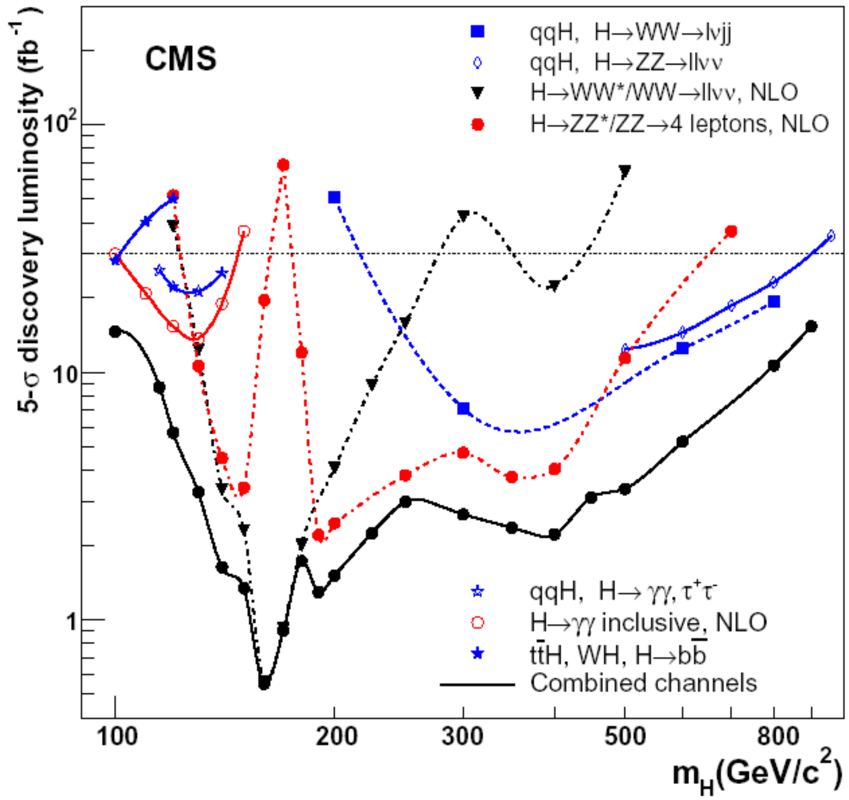
ATLAS

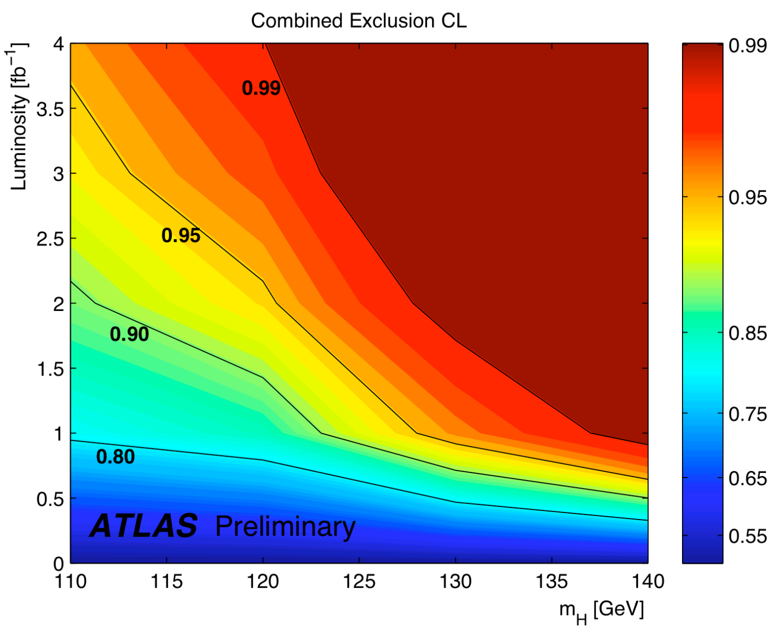


LHC has about 4 times better
signal / background than Tevatron

LHC SM Higgs Discovery Potential







- 5σ discovery over full mass range with ~20 fb^{-1}
  - Most challenging at low mass
- 95% exclusion over full mass range with ~4 fb^{-1}

Conclusions on Searches

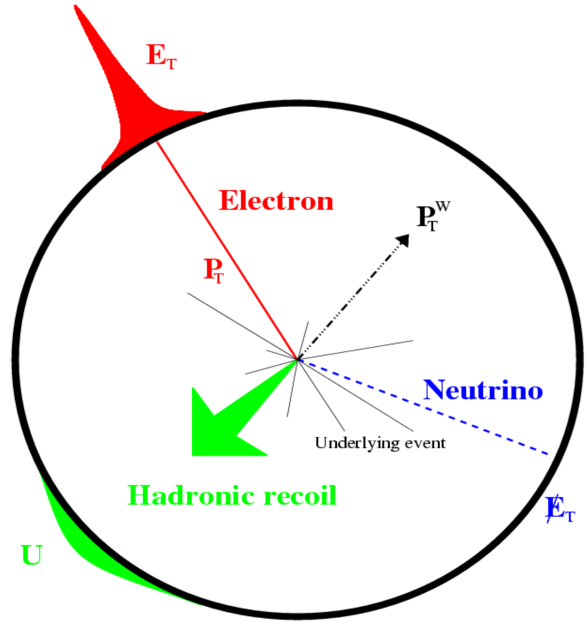
- Background estimate most crucial aspect for searches
- LHC has an amazing discovery potential
  - Supersymmetry already with ~100 pb^{-1}
    - Also other high mass particles, e.g.
    - Z’, Extra Dimensions, 4^{th} generation quarks, …
  - Higgs boson: 1-10 fb^{-1}
- Let’s hope that many exciting things will be found!!!

Measuring Properties of Particles

The $W^{\pm}$ Boson Mass

W Boson mass

- Real precision measurement:
  - LEP: M_W=80.367±0.033 GeV/c^2
  - Precision: 0.04%
    - => Very challenging!
- Main measurement ingredients:
  - Lepton p_T
  - Hadronic recoil parallel to lepton: $u_{||}$
- Z→ll superb calibration sample:
  - but statistically limited:
    - About a factor 10 less Z's than W's
    - Most systematic uncertainties are related to size of Z sample
      - Will scale with $1/\sqrt{N_Z}$ $(=1/\sqrt{L})$



$$m_T = \sqrt{2p_T^l \not{p}_T\,(1-\cos\Delta\phi)},$$

$$\not{p}_T \approx |p_T + u_{||}|$$

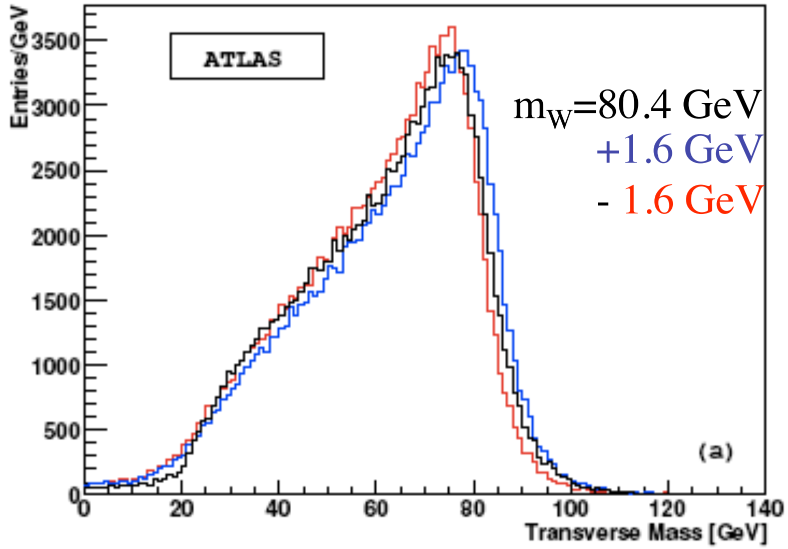
$$m_T \approx 2p_T\sqrt{1+u_{||}/p_T} \approx 2p_T + u_{||}$$

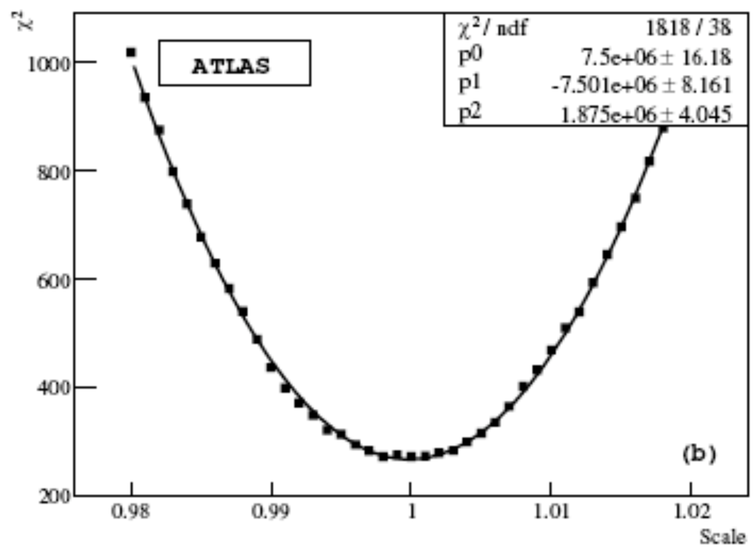
See arXiv:0708.3642

How to Extract the W Boson Mass

- Uses “Template Method”:
  - Templates created from MC simulation for different mW
  - Fit to determine which template fits best
  - Minimal $\chi^2 \Rightarrow$ W mass!
- Transverse mass of lepton and Met

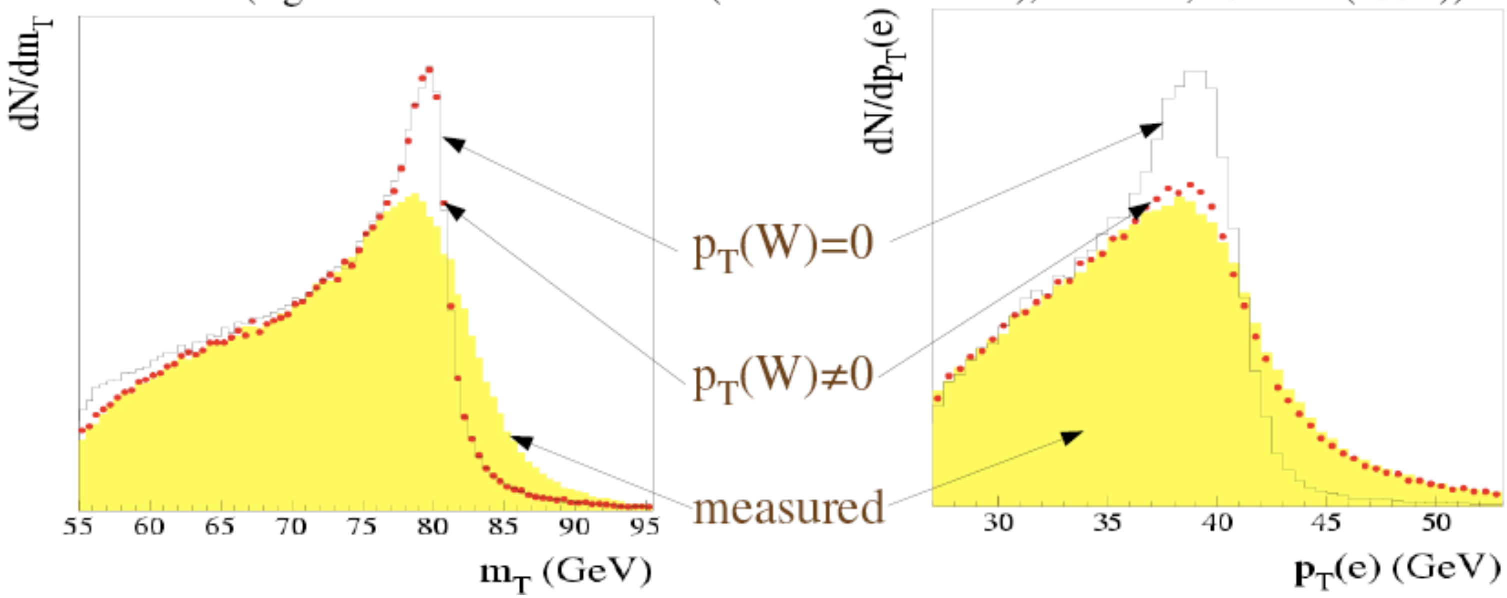
$$m_T = \sqrt{|p_T^\ell|^2 + |p_T^\nu|^2 - (\vec{p}_T^\ell + \vec{p}_T^\nu)^2}$$





How to Extract the W Boson Mass

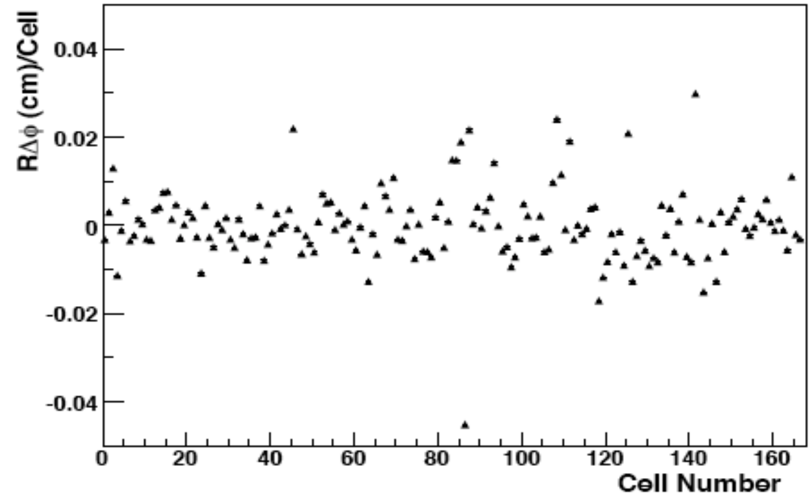
(figures from Abbott *et. al.* (D0 Collaboration), PRD 58, 092003 (1998))

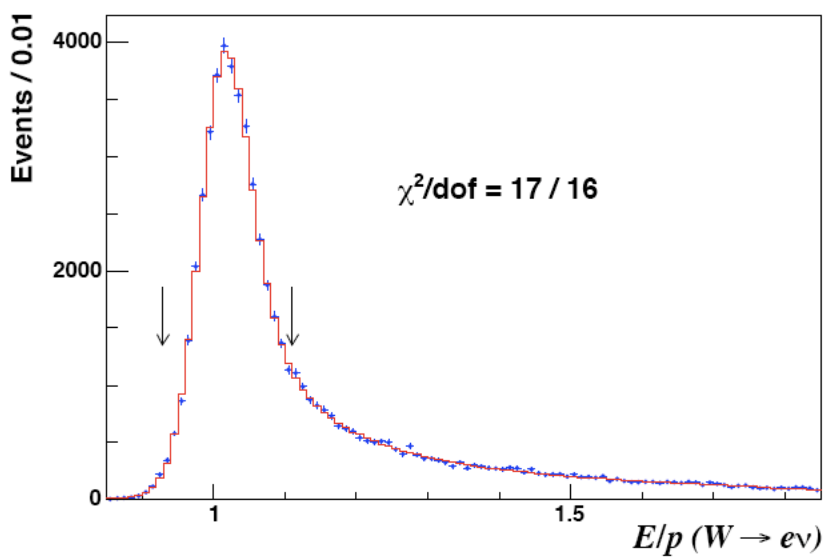


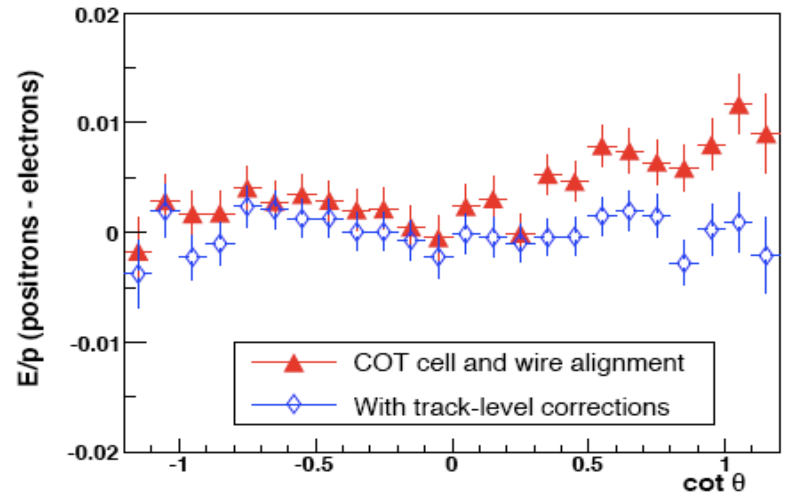
- Alternatively can fit to
  - Lepton p_T or missing E_T
- Sensitivity different to different systematics
  - Very powerful checks in this analysis:
    - Electrons vs muons
    - Z mass
    - m_T vs p_T vs ME_T fits
  - The redundancy is the strength of this difficult high precision analysis

Lepton Momentum Scale

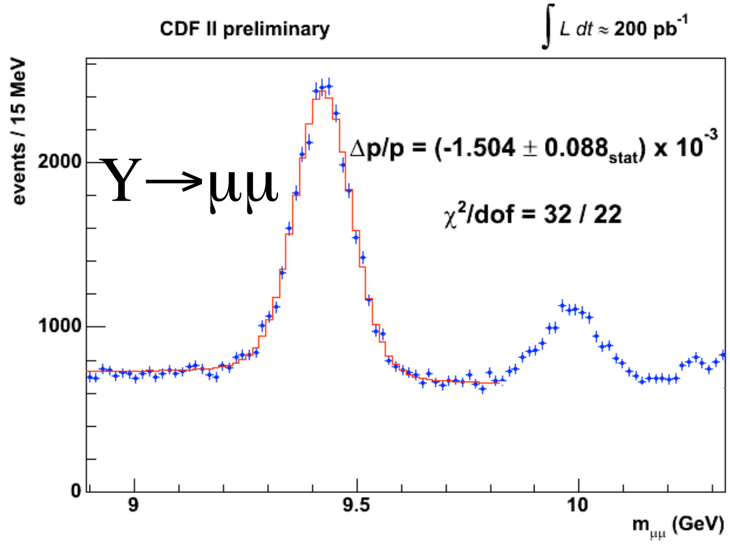
- Momentum scale:
  - Cosmic ray data used for detailed cell-by-cell calibration of CDF drift chamber
  - E/p of e+ and e- used to make further small corrections to p measurement
  - Peak position of overall E/p used to set electron energy scale
    - Tail sensitive to passive material

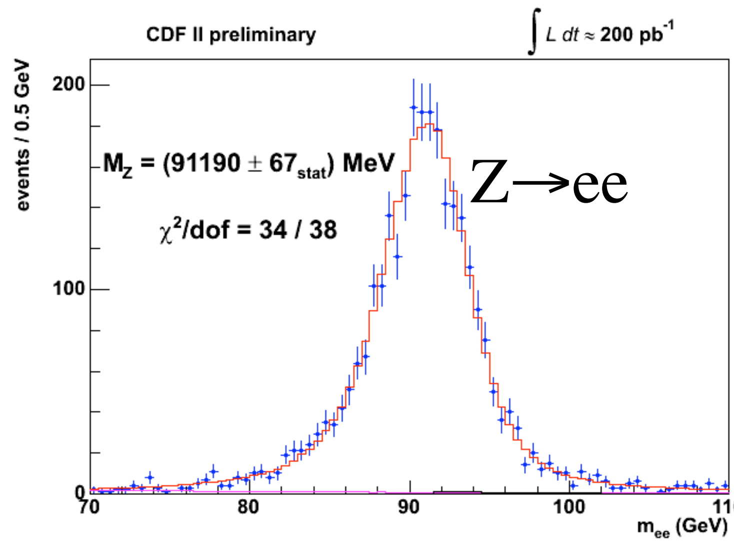


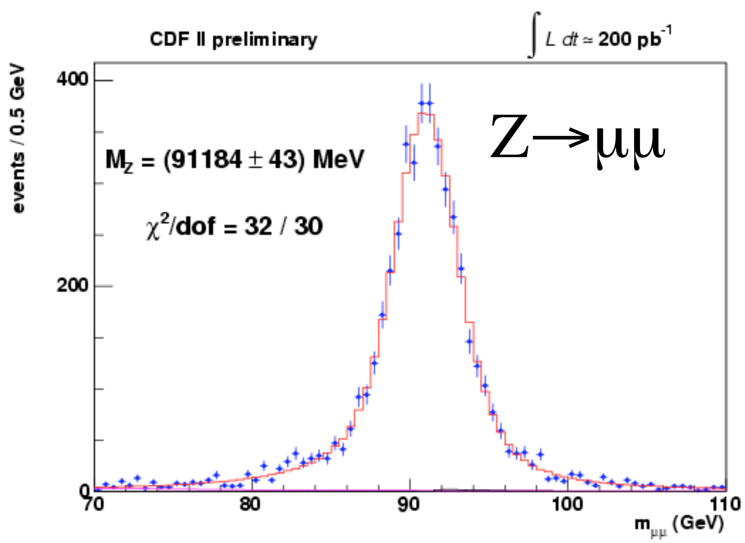


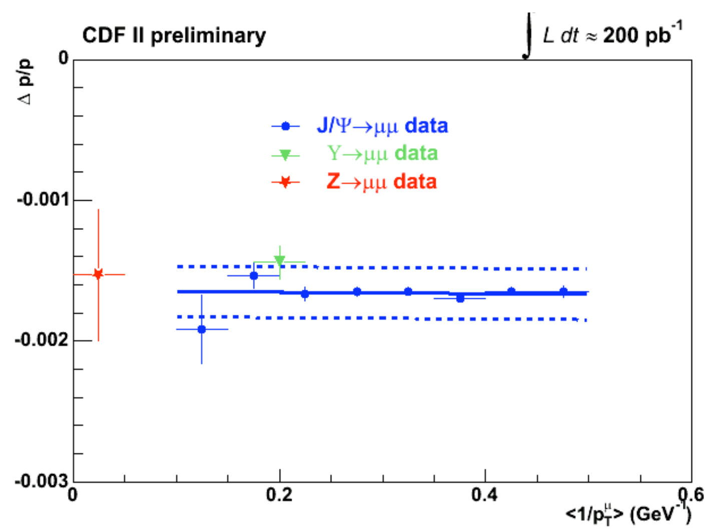


Momentum/Energy Scale and Resolution





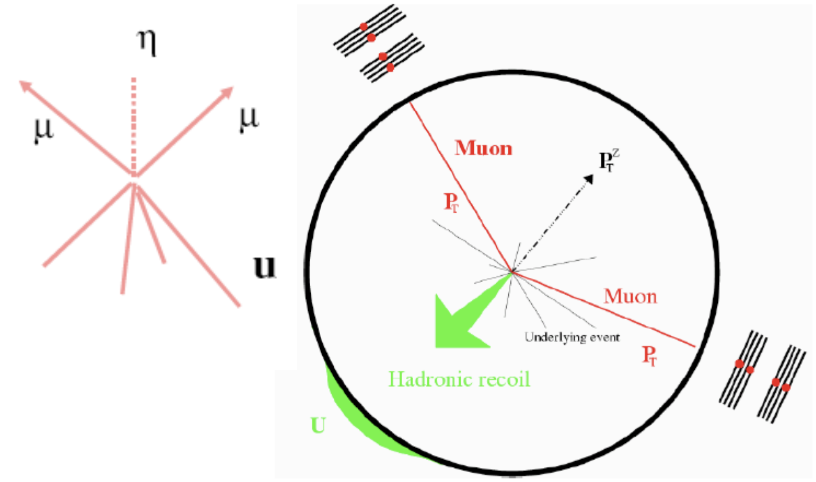


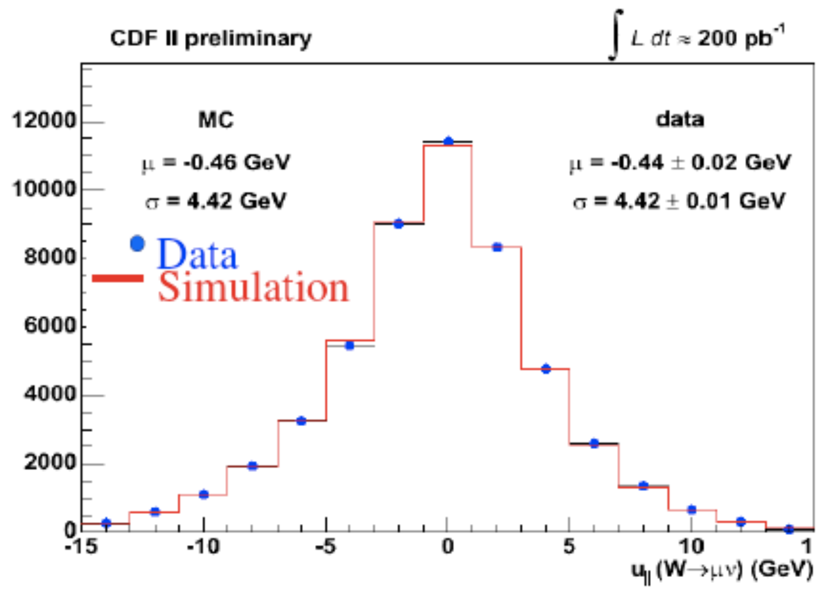


- Systematic uncertainty on momentum scale: 0.04%

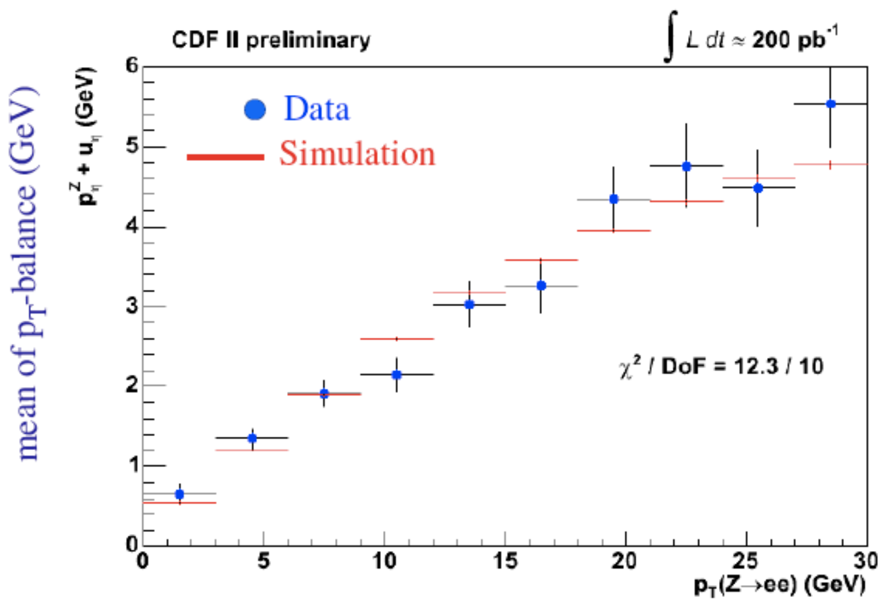
Hadronic Recoil Model

- Hadronic recoil modeling
  - Tune data based on Z's
  - Check on W's





Recoil projection (GeV) on lepton direction



Systematic Uncertainties

m_T Fit Uncertainties

Source	$W \to \mu\nu$	$W \to e\nu$	Correlation
Tracker Momentum Scale	17	17	100%
Calorimeter Energy Scale	0	25	0%
Lepton Resolution	3	9	0%
Lepton Efficiency	1	3	0%
Lepton Tower Removal	5	8	100%
Recoil Scale	9	9	100%
Recoil Resolution	7	7	100%
Backgrounds	9	8	0%
PDFs	11	11	100%
W Boson p_T	3	3	100%
Photon Radiation	12	11	100%
Statistical	54	48	0%
Total	60	62	-

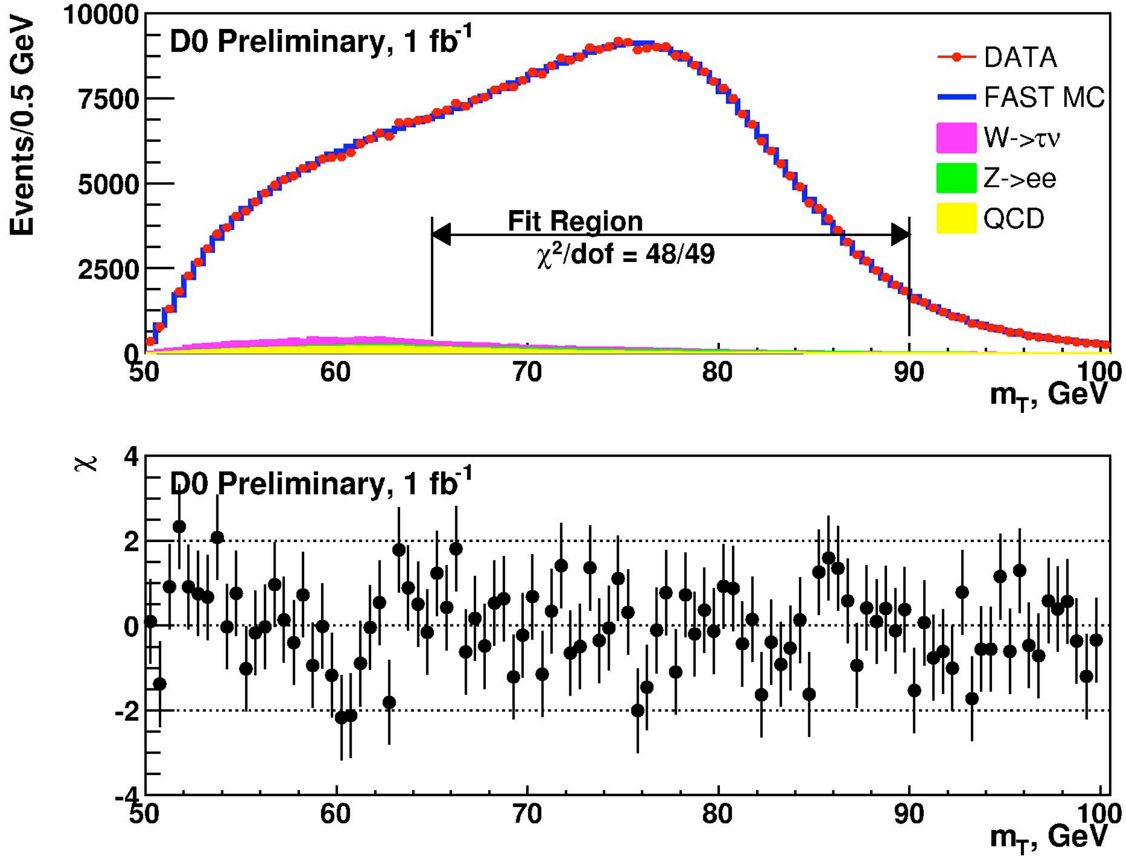
TABLE IX: Uncertainties in units of MeV on the transverse mass fit for m_W in the $W \to \mu\nu$ and $W \to e\nu$ samples.

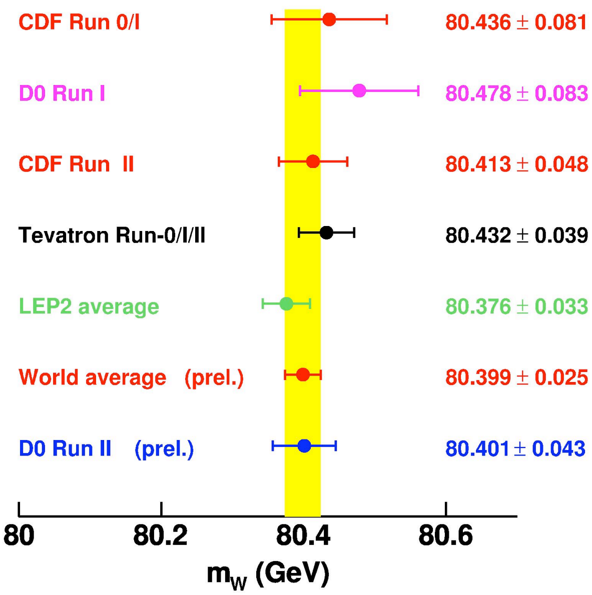
Limited by data statistics

Limited by data and theoretical understanding

- Overall uncertainty 60 MeV for both analyses
  - Careful treatment of correlations between them
- Dominated by stat. error (50 MeV) vs syst. (33 MeV)

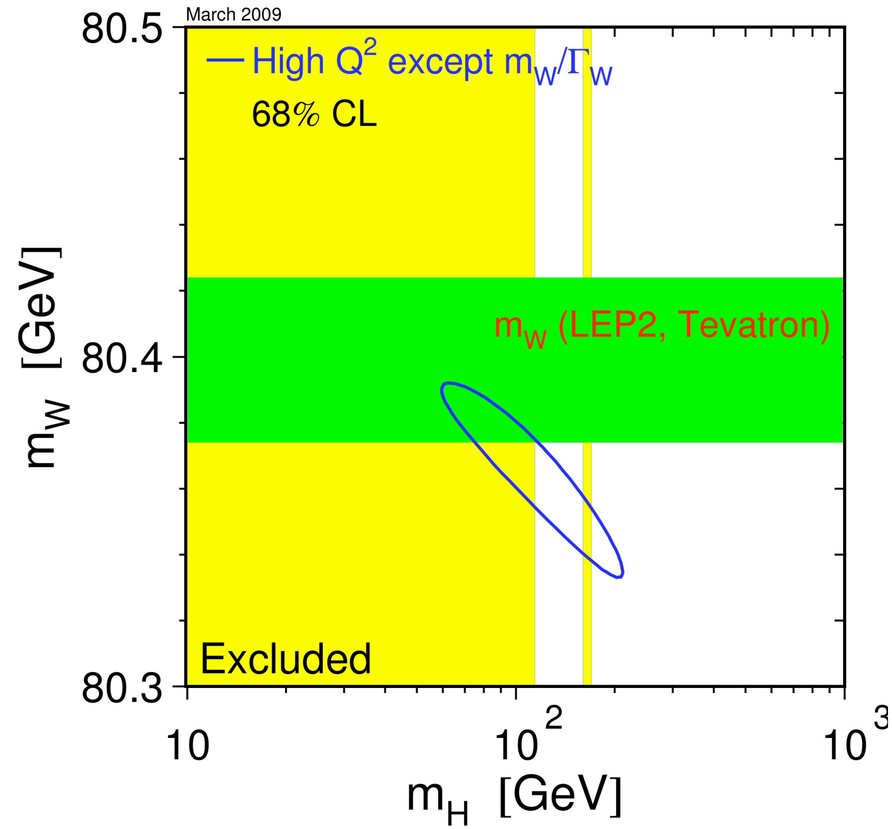
W Boson Mass Result

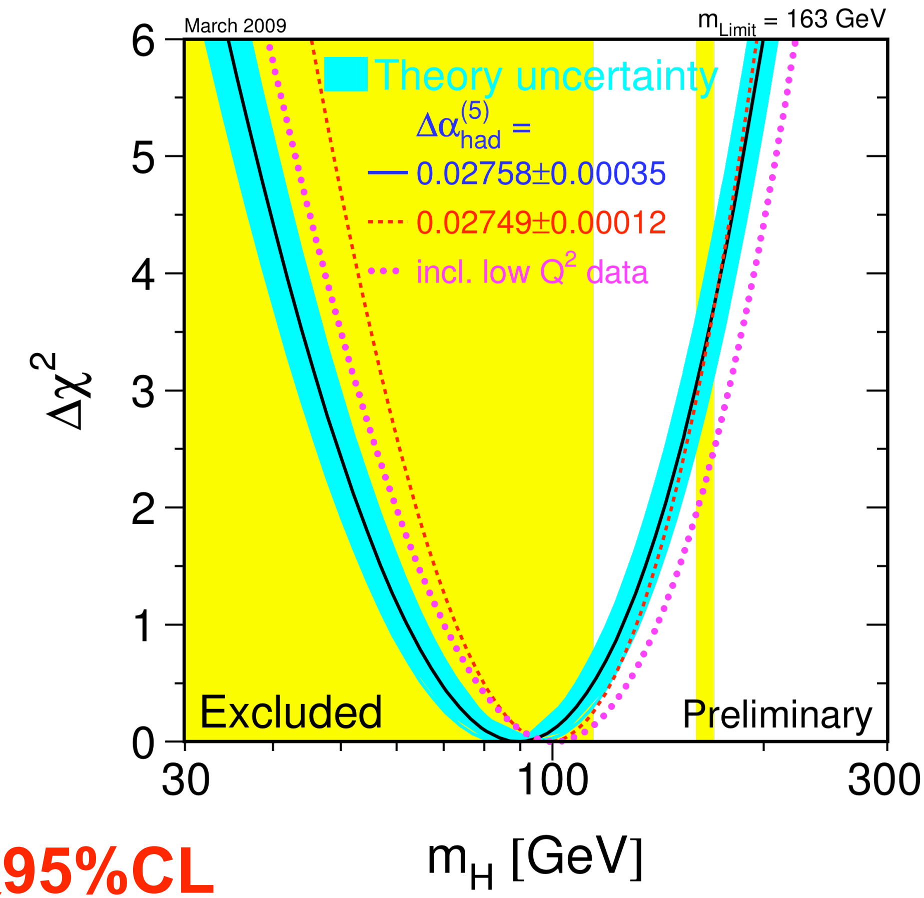




- New World average: **M_W=80399 ± 23 MeV**
- Ultimate Run 2 precision: **~15-20 MeV**

M_W, m_{top} and m_{Higgs}





- **Indirectly:** **m_H<163 GeV@95%CL**

(caveat: is the measured top mass the pole mass?)

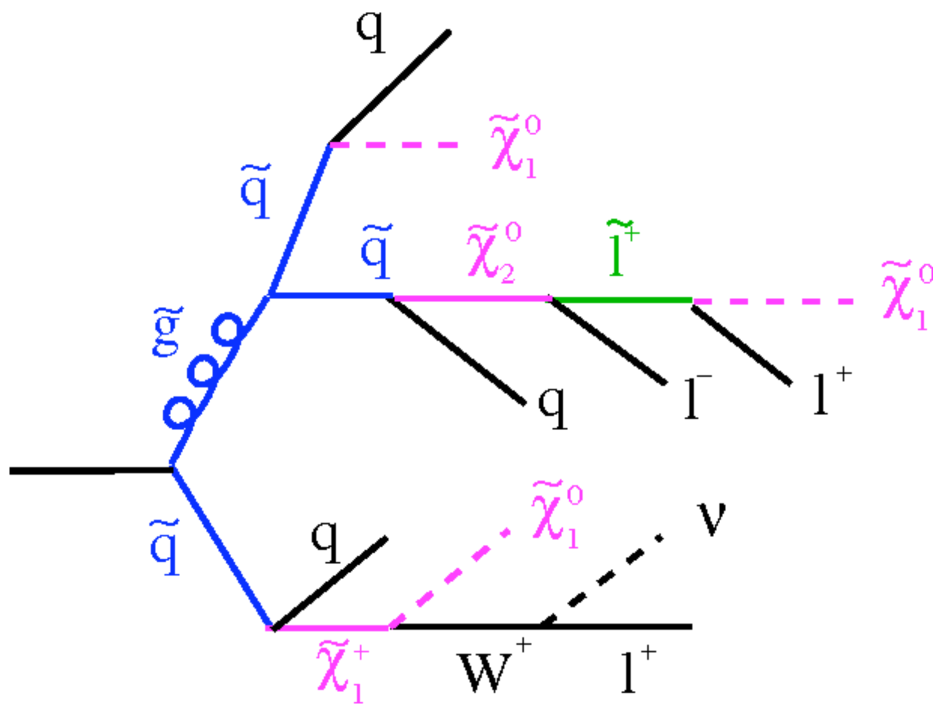
- **Directly:** **114<m_H<160 GeV or m_H>170 GeV @95%CL**

(This all assumes that there is no new physics beyond the SM)

Measuring Properties of Supersymmetric Particles (in case they exist)

Spectacular SUSY Events (?)

- Long cascade decays via several SUSY particles
  - In classic models quite possible
    - Would be a wonderful experimental challenge!
  - But of course very possible also that it does not happen



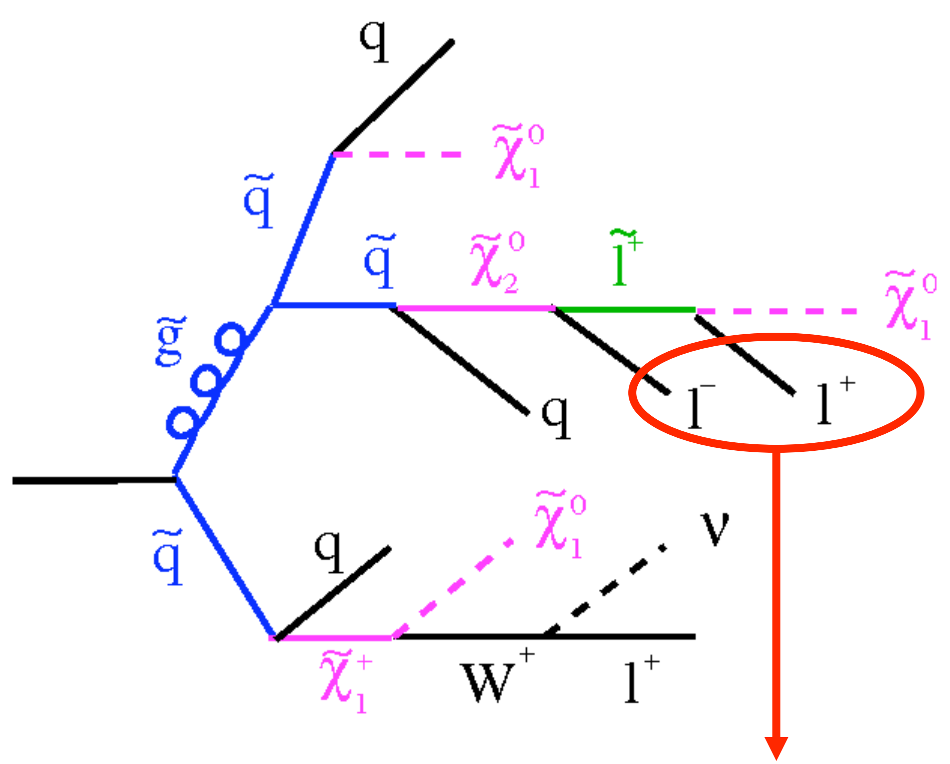
- If Nature is like this:
  - Need to try to reconstruct masses of all those particles
- Main method:
  - Measure “edges”

Spectacular SUSY Events (?)

- Long cascade decays via several SUSY particles, e.g.

$$\boxed{\tilde{q}_L \rightarrow \tilde{\chi}_2^0 q (\rightarrow \tilde{\ell}^{\pm} \ell^{\mp} q) \rightarrow \tilde{\chi}_1^0 \ell^+ \ell^- q}$$

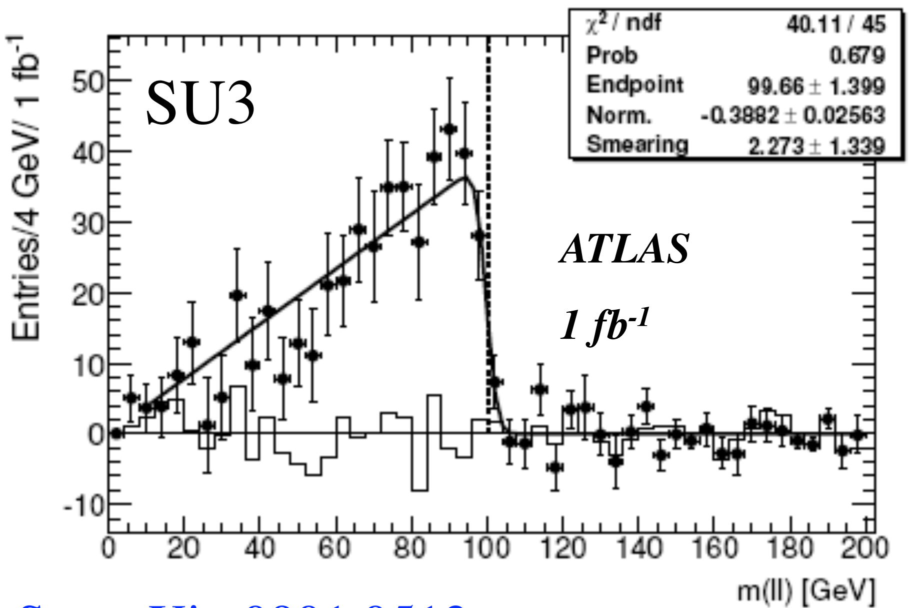
  - In classic models quite possible
    - Would be a wonderful experimental challenge!
  - But of course very possible also that it does not happen
- If Nature is like this:
  - Need to try to reconstruct masses of all those particles
- Main method:
  - Measure "edges"

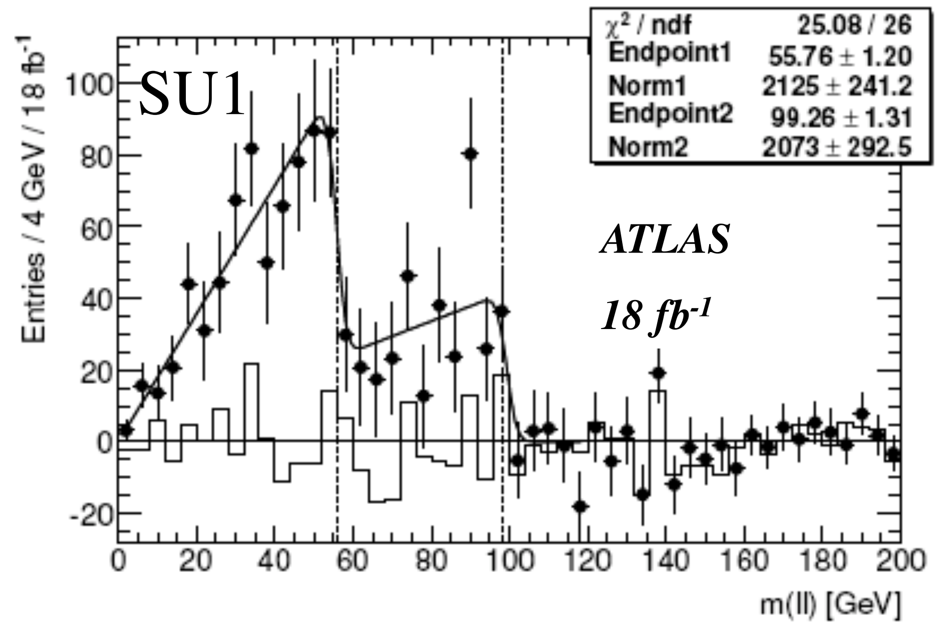


$$\boxed{m_{\ell\ell}^{\mathrm{edge}} = m_{\tilde{\chi}_2^0} \sqrt{1 - \left(\frac{m_{\tilde{\ell}}}{m_{\tilde{\chi}_2^0}}\right)^2} \sqrt{1 - \left(\frac{m_{\tilde{\chi}_1^0}}{m_{\tilde{\ell}}}\right)^2}\,.}$$

Only for opposite sign same-flavor (OSDF) leptons

Dilepton Edge Fit





See arXiv:0901.0512

- Background from different flavors subtracted $\Sigma e^+e^- + \mu^+\mu^- - e^+\mu^- - \mu^+e^-$
  - Removes random SUSY backgrounds, top backgrounds,..
- Fit for dilepton edge
  - With many such edges one can maybe get a beginning of an understanding what is happening!
  - Different models look differently

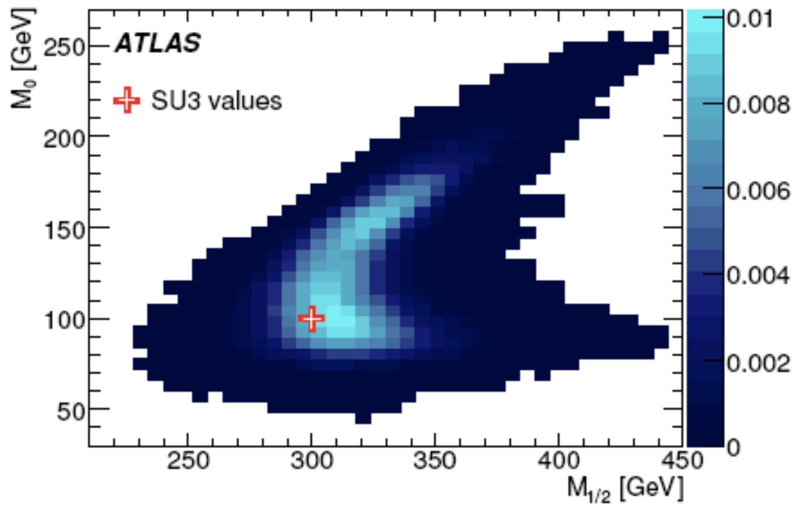
$$m_{\ell\ell}^{\text{edge}} = m_{\tilde{\chi}_2^0}\sqrt{1-\left(\frac{m_{\tilde{\ell}}}{m_{\tilde{\chi}_2^0}}\right)^2}\sqrt{1-\left(\frac{m_{\tilde{\chi}_1^0}}{m_{\tilde{\ell}}}\right)^2}.$$

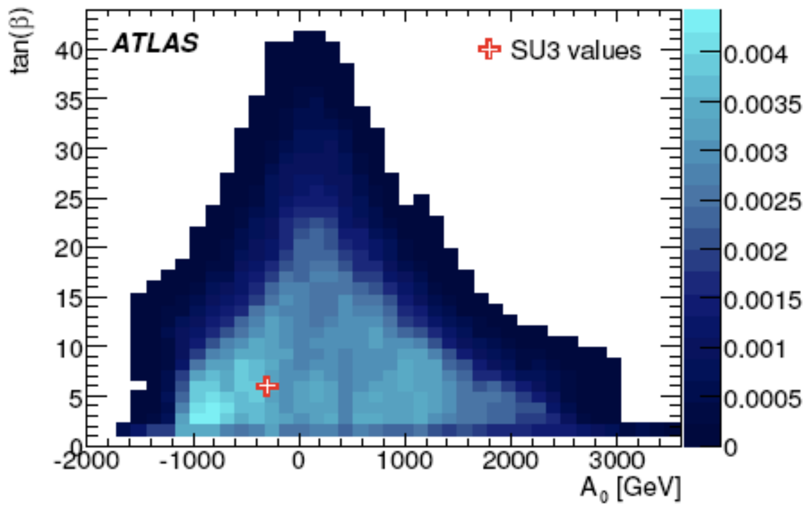
How well does this work?

Endpoint	SU3 truth	SU3 measured	SU4 truth	SU4 measured
$m_{\ell\ell q}^{\text{edge}}$	501	$517 \pm 30 \pm 10 \pm 13$	340	$343 \pm 12 \pm 3 \pm 9$
$m_{\ell\ell q}^{\text{thr}}$	249	$265 \pm 17 \pm 15 \pm 7$	168	$161 \pm 36 \pm 20 \pm 4$
$m_{lq(\text{low})}^{\text{max}}$	325	$333 \pm 6 \pm 6 \pm 8$	240	$201 \pm 9 \pm 3 \pm 5$
$m_{lq(\text{high})}^{\text{max}}$	418	$445 \pm 11 \pm 11 \pm 11$	340	$320 \pm 8 \pm 3 \pm 8$

- Works reasonably well…
- Can even try to extract high-level theory parameters

SUSY Parameters at GUT scale!?!





Parameter	SU3 value	fitted value	exp. unc.
		$\mathrm{sign}(\mu) = +1$	
$\tan\beta$	6	7.4	4.6
M_0	100 GeV	98.5 GeV	± 9.3 GeV
$M_{1/2}$	300 GeV	317.7 GeV	± 6.9 GeV
A_0	-300 GeV	445 GeV	± 408 GeV
		$\mathrm{sign}(\mu) = -1$	
$\tan\beta$		13.9	± 2.8
M_0		104 GeV	± 18 GeV
$M_{1/2}$		309.6 GeV	± 5.9 GeV
A_0		489 GeV	± 189 GeV

- Depends if we understand our model well enough
- Personally I am very skeptical that we can do this
  - But would be great to have that problem!

Conclusions on Measuring Properties

- Several methods of extracting property of particle
  - Template method is widely used
  - Matrix Element technique extracts more information
  - For known shapes simple fits can also be done
- Examples:
  - W boson mass (precision ~0.06%)
  - Top quark mass (precision ~0.7%)
  - SUSY particles masses (precision ~unknown)
- Critical to understand detector calibration
  - Utilize known resonances
- I hope we will be able to measure properties of many new particles

Concluding Remarks

Data are very precious

- Treat them with the highest respect
- Try to not jump to conclusions too fast
  - Data analysis is like detective work
- Try to use all you can to understand them
  - Redundancy of detector
    - tracker vs calorimeter etc.
  - Complementary physics processes
    - W's vs Z's etc.
  - Monte Carlo tools and theoretical calculations
- Above all: use your brain and your judgment

This was my very personal view on the key issues of data analysis

Thanks,

and lot's of fun and luck for your analyses

LHC Expectations: W mass

See arXiv:0901.0512

Method	$p_T(e)$ [MeV]	$p_T(\mu)$ [MeV]	$M_T(e)$ [MeV]	$M_T(\mu)$ [MeV]
δm_W (stat)	120	106	61	57
δm_W (α_E)	110	110	110	110
δm_W (σ_E)	5	5	5	5
δm_W (tails)	28	< 28	28	< 28
δm_W (ε)	14	–	14	–
δm_W (recoil)	–	–	200	200
δm_W (bkg)	3	3	3	3
δm_W (exp)	114	114	230	230
δm_W (PDF)	25	25	25	25
Total	167	158	239	238

- Expect uncertainty of 150-250 MeV with 15 pb^{-1}
  - Ultimately expected to improve upon Tevatron precision when detector well understood…

Measuring an Asymmetry

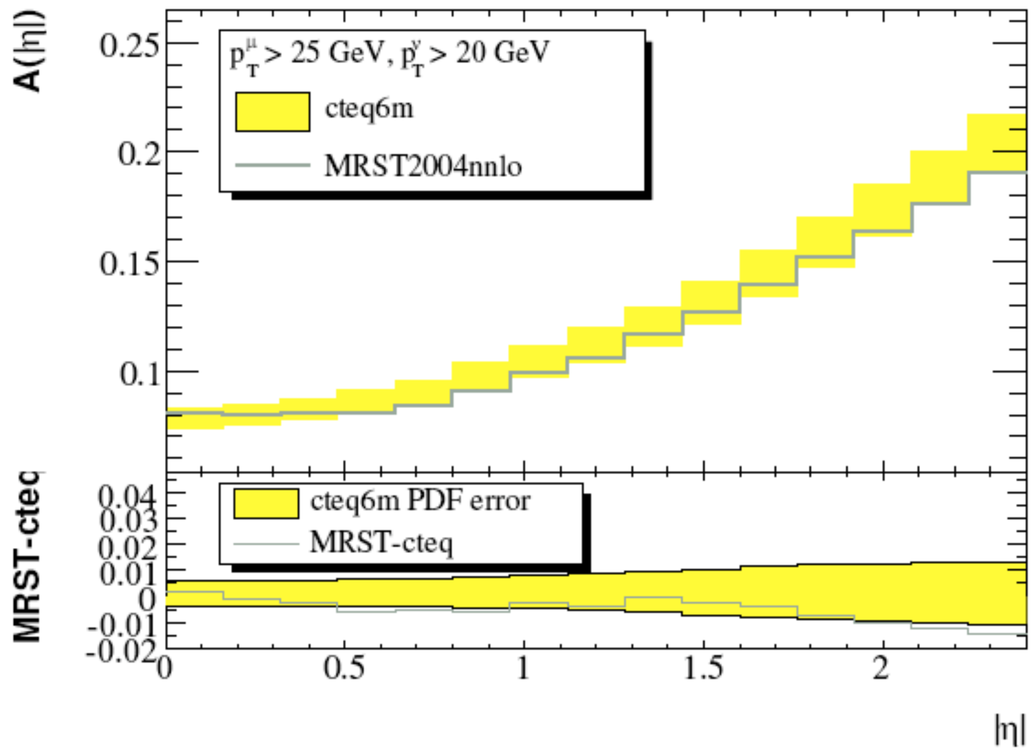
Asymmetries

- Many important asymmetries have been measured, e.g.
  - Forward-backward asymmetry in Z's: A_{FB}
    - sensitive to photon/Z interference and new physics
  - W+/W- charge asymmetry:
    - Sensitive to parton distribution functions
  - B meson decay asymmetries
    - Sensitive to matter/anti-matter differences
  - …
- Experimental advantage:
  - Many systematic uncertainties (partially) cancel

W Boson Charge Asymmetry: LHC

- W charge asymmetry arises due to more up-quarks compared to down-quarks in proton
  - $N(W^+)/N(W^-) \approx 1.5$
    - For $\sqrt{s}$=10 TeV
  - Depends on $|\eta|$
  - Sensitive to ratios of u and d-quarks densities

$$A(\eta) = \frac{\frac{d\sigma}{d\eta}(W^+ \to \mu^+\bar{\nu}_\mu) - \frac{d\sigma}{d\eta}(W^- \to \mu^-\nu_\mu)}{\frac{d\sigma}{d\eta}(W^+ \to \mu^+\bar{\nu}_\mu) + \frac{d\sigma}{d\eta}(W^- \to \mu^-\nu_\mu)}$$



See EWK-09-003 and EWK-08-002 at https://twiki.cern.ch/twiki/bin/view/CMS/PhysicsResults

Event Selection

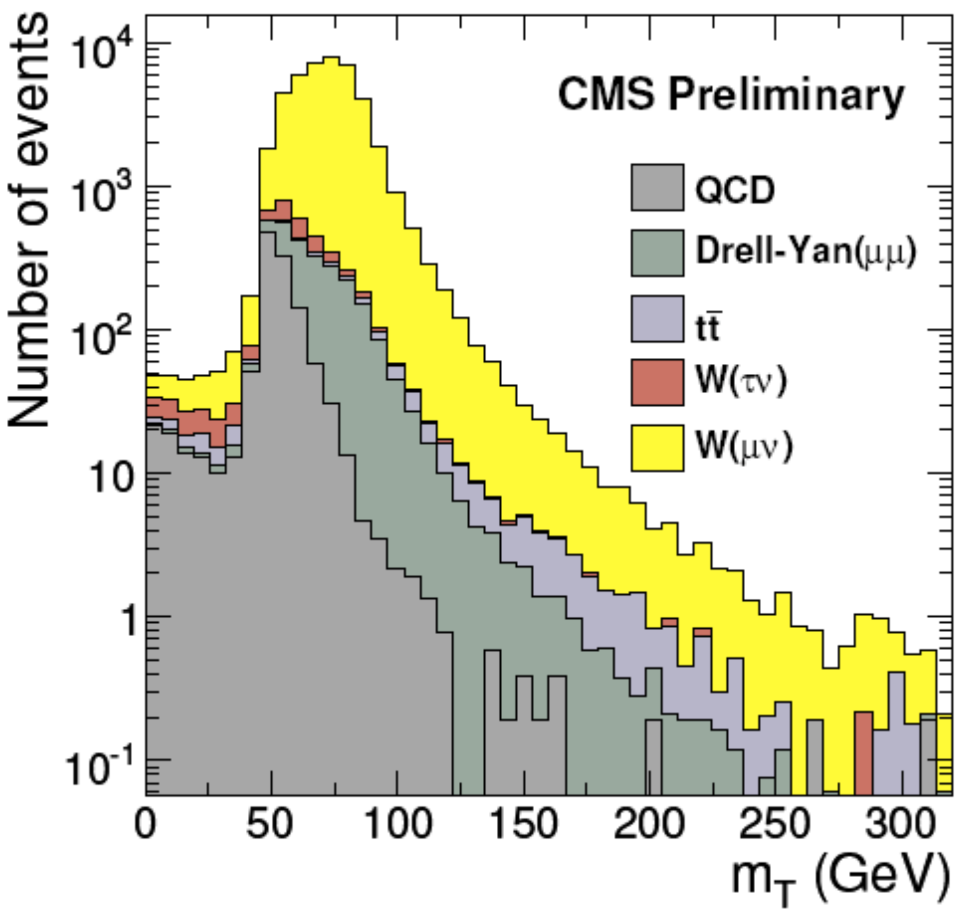
Selection	Efficiency (%) $W^+ \to \mu^+\nu$	$W^- \to \mu^-\nu$
One loose HLT-matched Muon ($\|\eta\|$ <2.1)	77.1±0.1	81.8±0.1
Muon $p_T + Iso$ >25 GeV	77.8±0.1	82.4±0.1
z <0.05	92.4±0.1	92.3±0.1
MET > 20 GeV	96.3±0.1	97.3±0.1
Total Efficiency	53.4±0.1	60.6±0.1

$$z = 1 - \frac{p_T}{p_T + Iso}.$$

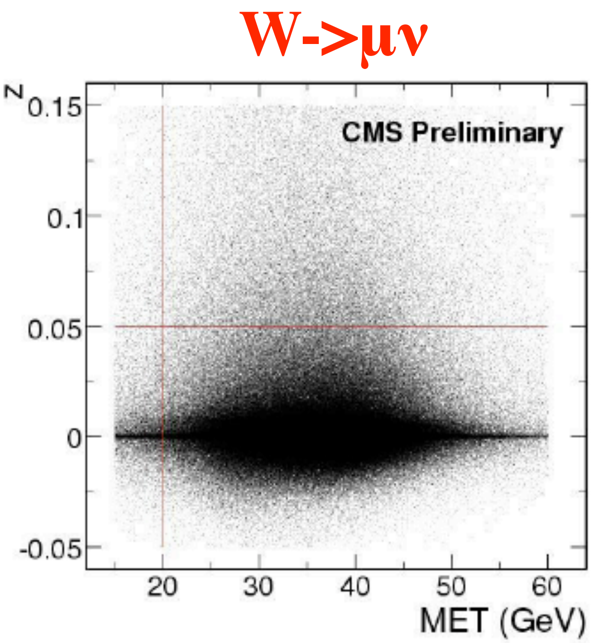
- Note:
  - Different efficiencies for W^+ and W^- events (~10%)
  - Due to different kinematic acceptance
    - p_T and η cuts

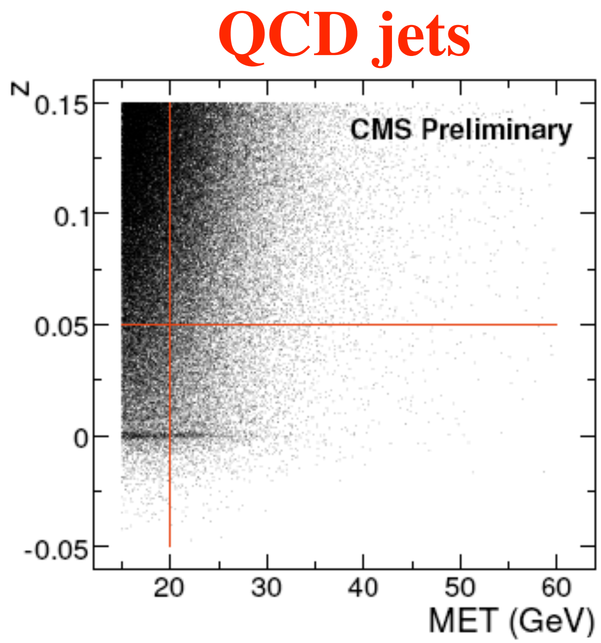
Backgrounds

- Background will generally have different charge asymmetry than signal
- Typical background:
  - ~10%
  - Need to determine asymmetry for backgrounds
    - Define orthogonal unbiased selection to e.g. measure it for QCD jet background

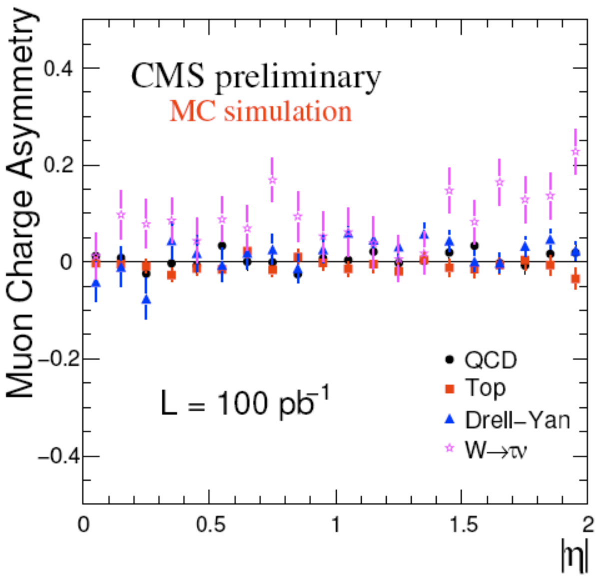


QCD Background





- E.g. use region of poor isolation and low MET to develop background model
  - E.g. z>0.05 and Met<20 GeV
- Measure asymmetry for those events
  - Vary method to assess systematics



From Counting Events to A(η)

We measure:

$$A(\eta)_{obs} = \frac{\frac{dN^+}{d\eta} - \frac{dN^-}{d\eta}}{\frac{dN^+}{d\eta} + \frac{dN^-}{d\eta}},$$

(after background subtraction)

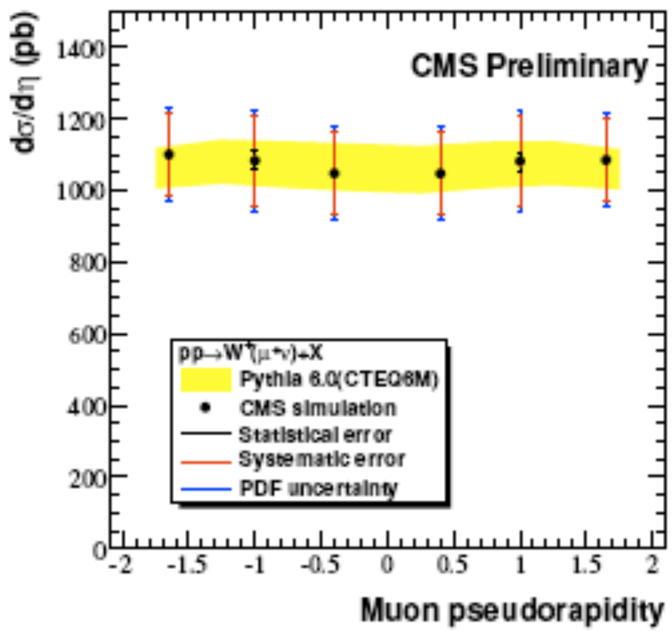
Related to cross section: $$\frac{dN}{d\eta} = \mathcal{L} \cdot \frac{d\sigma}{d\eta} \cdot \epsilon_{HLT} \cdot \epsilon_{offline} \cdot \epsilon_{acceptance}$$

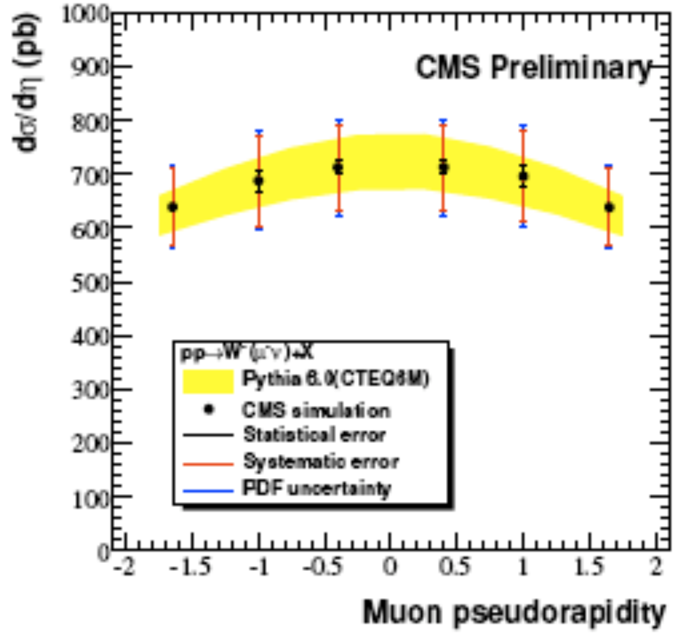
Finally we get:

$$A(\eta) = \frac{\frac{dN^+}{d\eta} - \frac{dN^-}{d\eta} \cdot \frac{\epsilon^+_{HLT} \cdot \epsilon^+_{offline} \cdot \epsilon^+_{acceptance}}{\epsilon^-_{HLT} \cdot \epsilon^-_{offline} \cdot \epsilon^-_{acceptance}}}{\frac{dN^+}{d\eta} + \frac{dN^-}{d\eta} \cdot \frac{\epsilon^+_{HLT} \cdot \epsilon^+_{offline} \cdot \epsilon^+_{acceptance}}{\epsilon^-_{HLT} \cdot \epsilon^-_{offline} \cdot \epsilon^-_{acceptance}}}$$

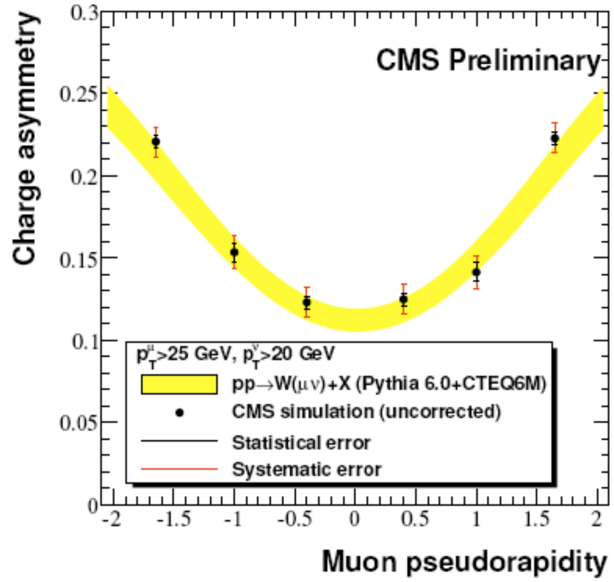
- Note:
  - Need to know efficiencies as function of η
  - Neglected charge misidentification here

Cross Section vs η



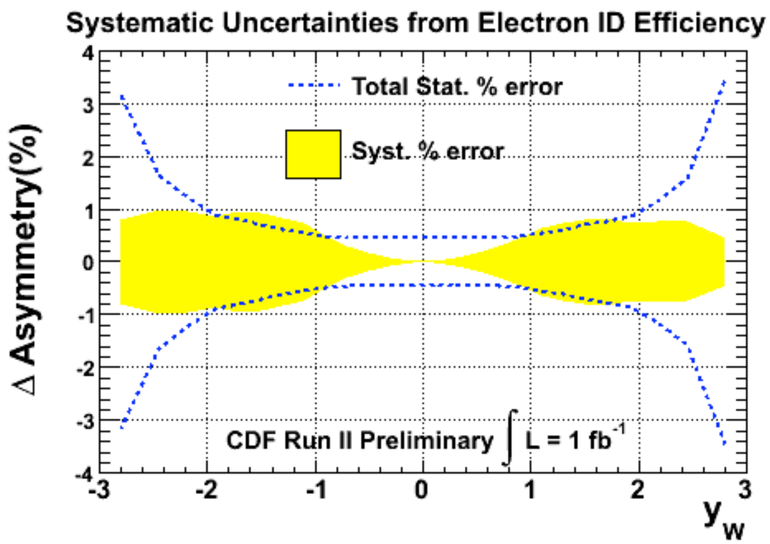


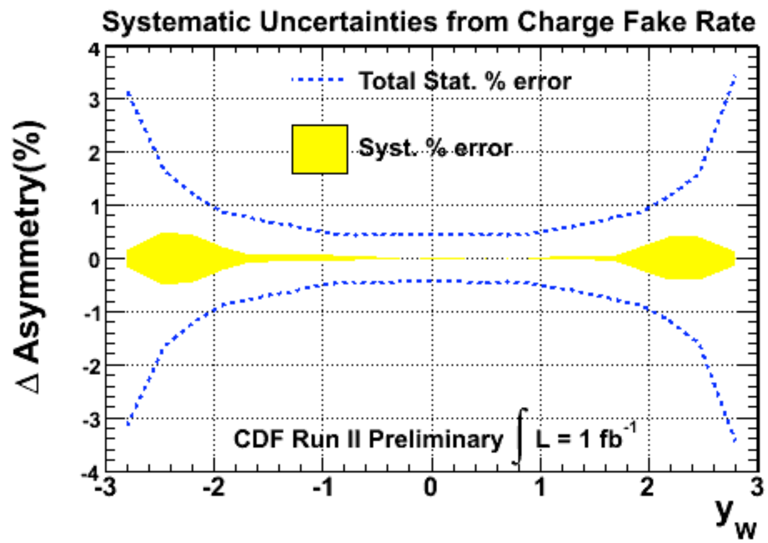
- Precision as expected for 10 pb^{-1} of LHC data
- Experimental errors comparable to theory uncertainties



Systematic Uncertainties

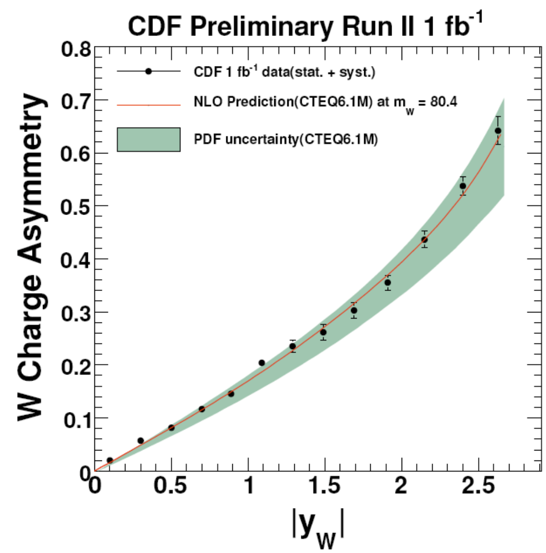
- Examples from similar measurement of CDF





In general similar to cross section Measurement

- beware that not everything cancels
- currently D0 and CDF disagree

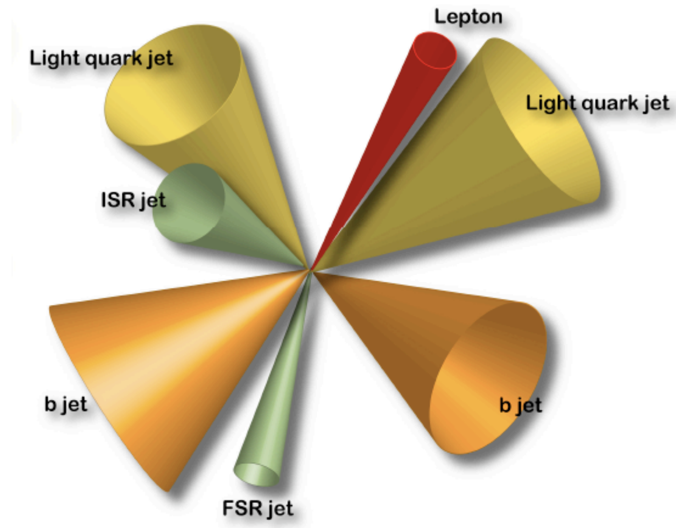


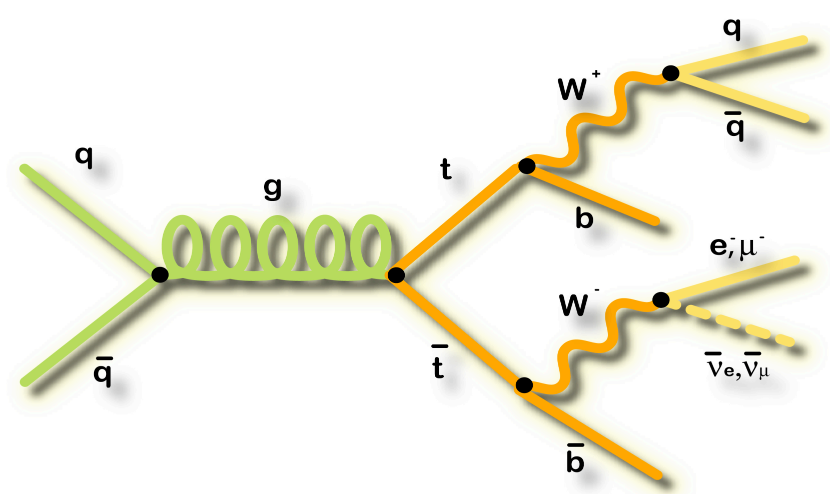
See http://www-cdf.fnal.gov/physics/ewk/2007/WChargeAsym/W_Charge_Asymmetry.html

The Top Quark Mass

Top Mass Measurement: tt→(blν)(bqq)

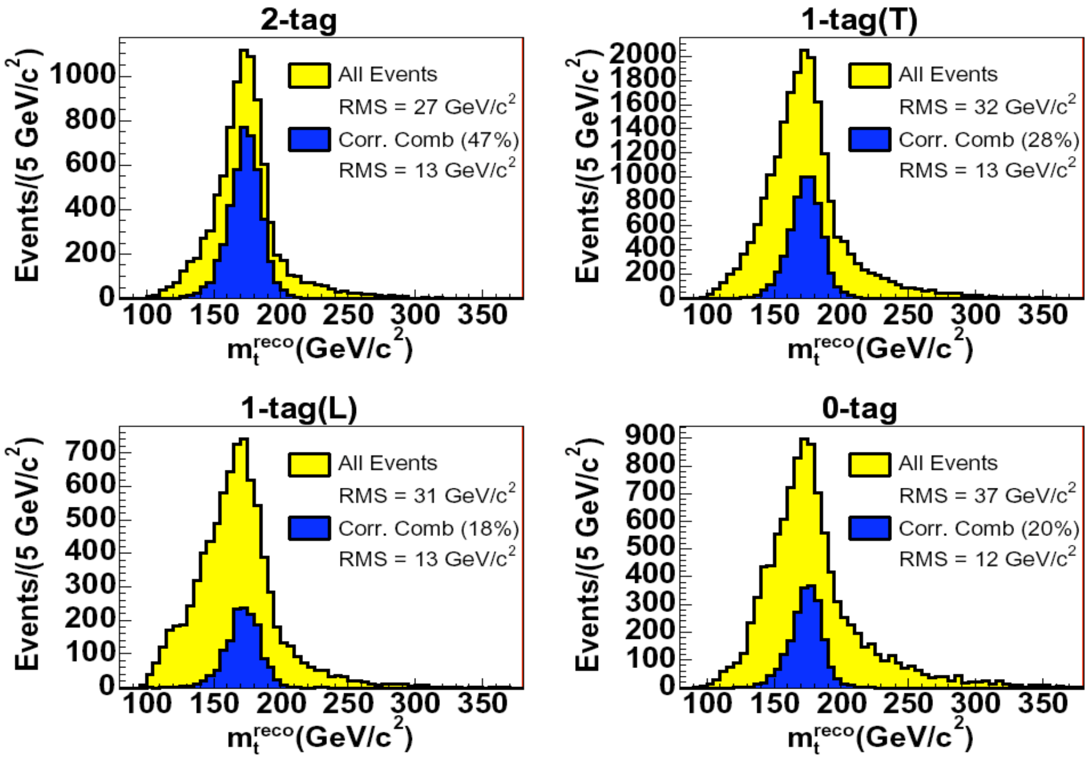
- 4 jets, 1 lepton and missing E_T
  - Which jet belongs to what?
  - Combinatorics!
- B-tagging helps:
  - 2 b-tags =>2 combinations
  - 1 b-tag => 6 combinations
  - 0 b-tags =>12 combinations
- Two Strategies:
  - Template method: –
    - Uses “best” combination
    - Chi2 fit requires m(t)=m(t)
  - Matrix Element method:
    - Uses all combinations
    - Assign probability depending on kinematic consistency with top





Top Mass Determination

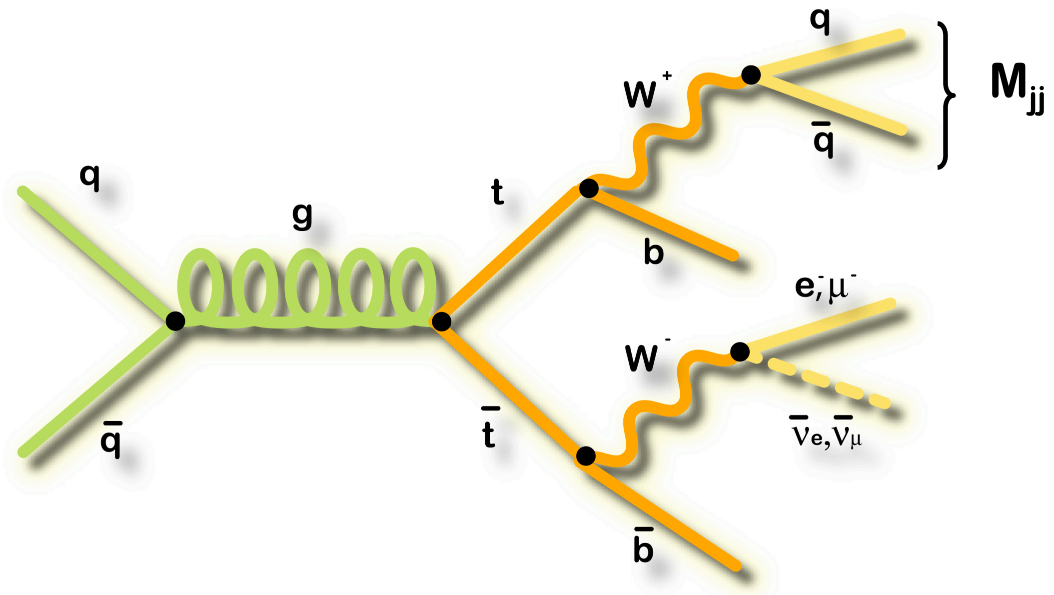
- Inputs:
  - Jet 4-vectors
  - Lepton 4-vector
  - Remaining transverse energy, $p_{T,UE}$:
    - $p_{T,\nu}=-(p_{T,l}+p_{T,UE}+\sum p_{T,jet})$
- Constraints:
  - $M(l\nu)=M_W$
  - $M(q\bar{q})=M_W$
  - $M(t)=M(\bar{t})$
- Unknown:
  - Neutrino p_z
- 1 unknown, 3 constraints:
  - Overconstrained
  - Can measure M(t) for each event: m_t^{reco}



Selecting correct combination 20-50% of the time

In-situ Measurement of JES

- Additionally, use W→jj mass resonance (M_{jj}) to measure the jet energy scale (JES) uncertainty

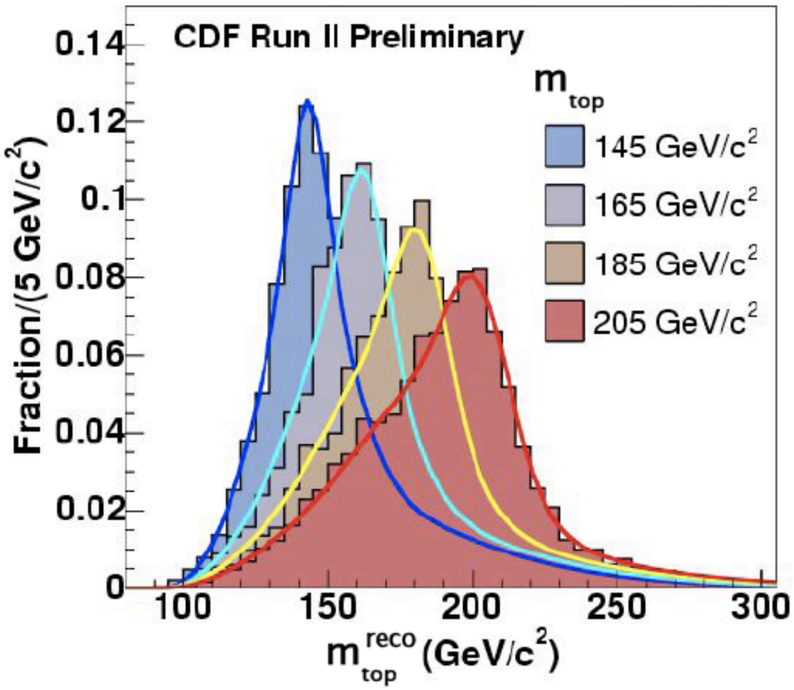


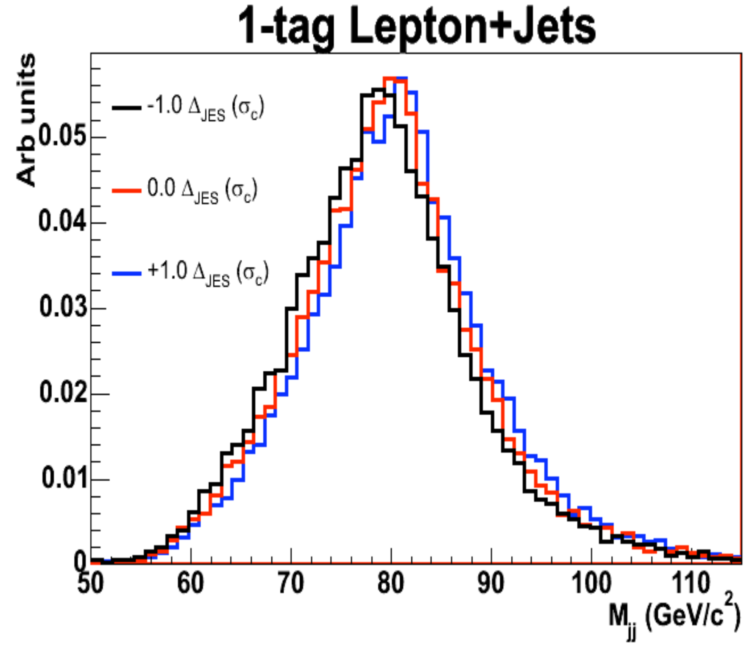
2D fit of the invariant mass of the non-b-jets and the top mass:

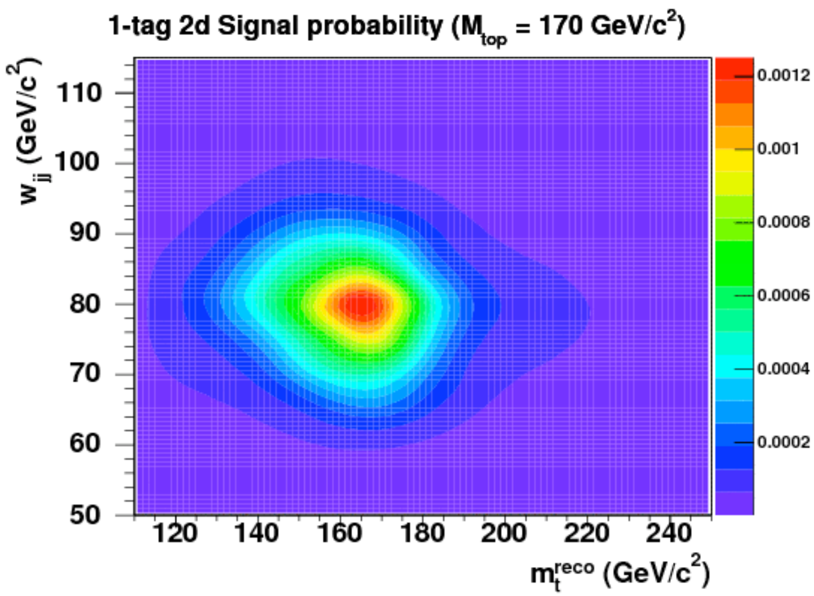
JES $\propto$ M(jj)- 80.4 GeV/c^2

Measurement of JES scales directly with data statistics

Top Mass Templates



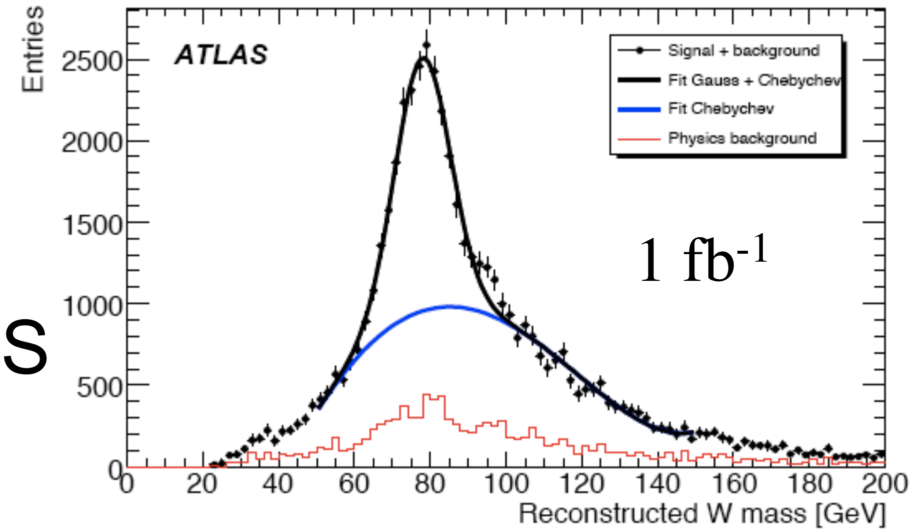


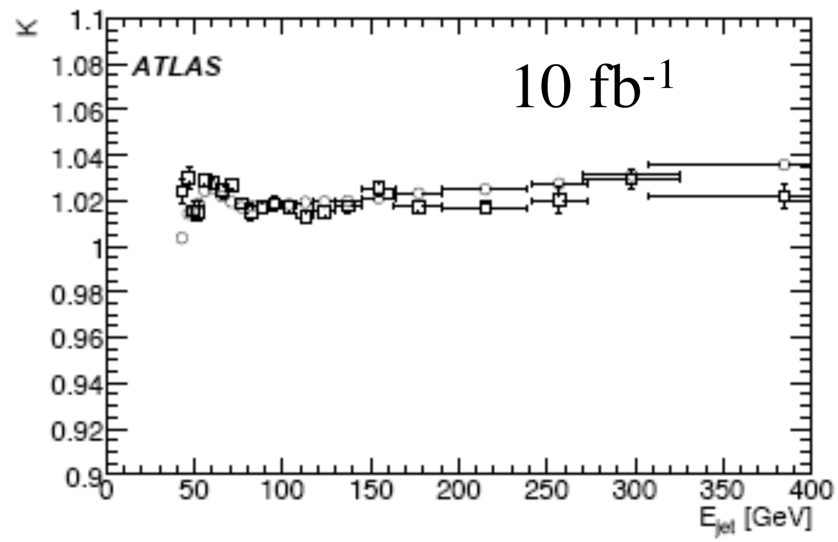


- Fit to those templates for
  - Top mass
  - Jet Energy Scale

Measurement of JES at LHC

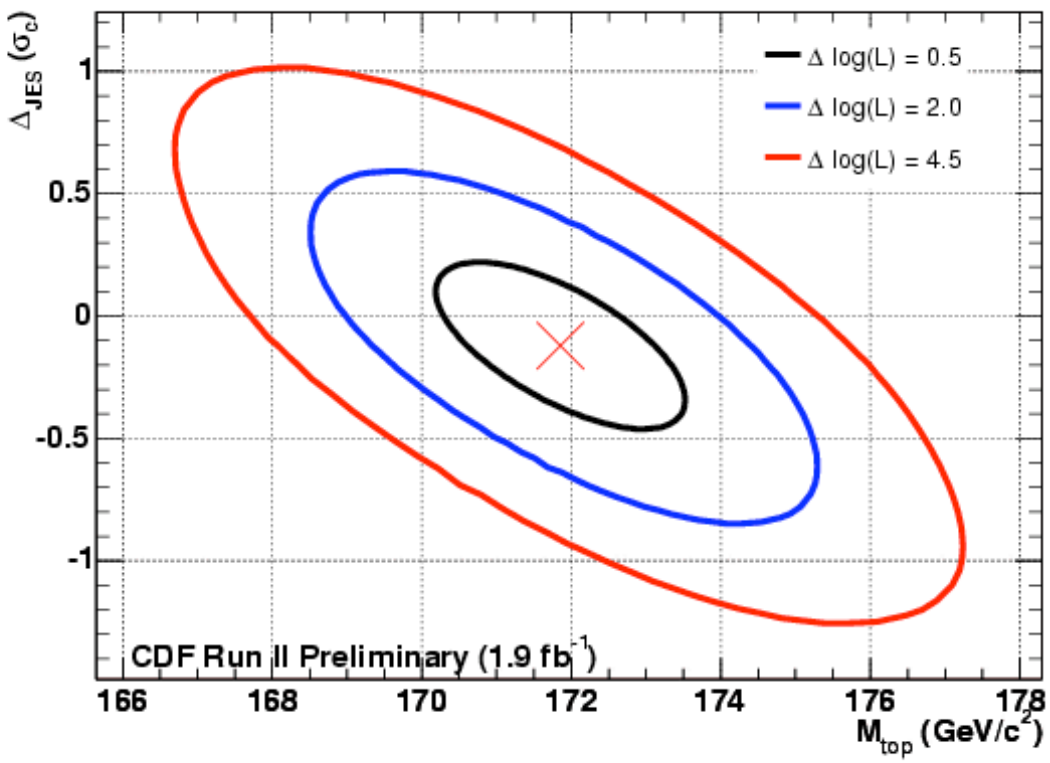
- Large top samples
  - Clean W mass peak
- Allow measurement of JES as function of Jet Energy
- Can achieve 1% precision with 10 fb^{-1}





Template Analysis Results on m_{top}

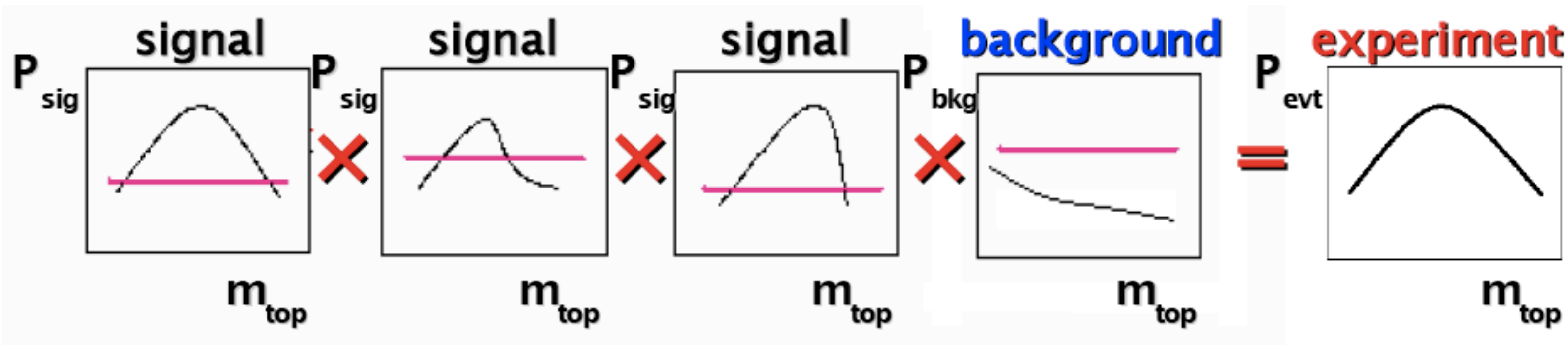
- Using 344 lepton+jets and 144 dilepton candidate events in 1.9 fb^{-1}
- Using in-situ JES calibration results in factor four improvement on JES



$$m_{top} = 171.9 \pm 1.7\ (\text{stat.+JES}) \pm 1.0 = 171.6 \pm 2.0\ \text{GeV/c}^2$$

"Matrix Element Method"

- Construct probability density function as function of m_{top} for each event
- Multiply those probabilities of all events

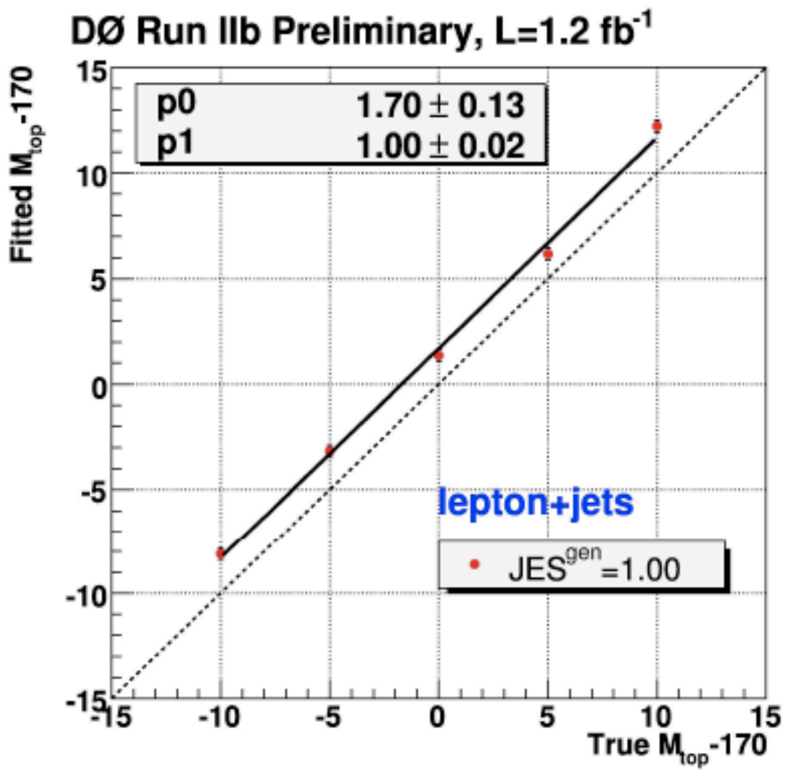


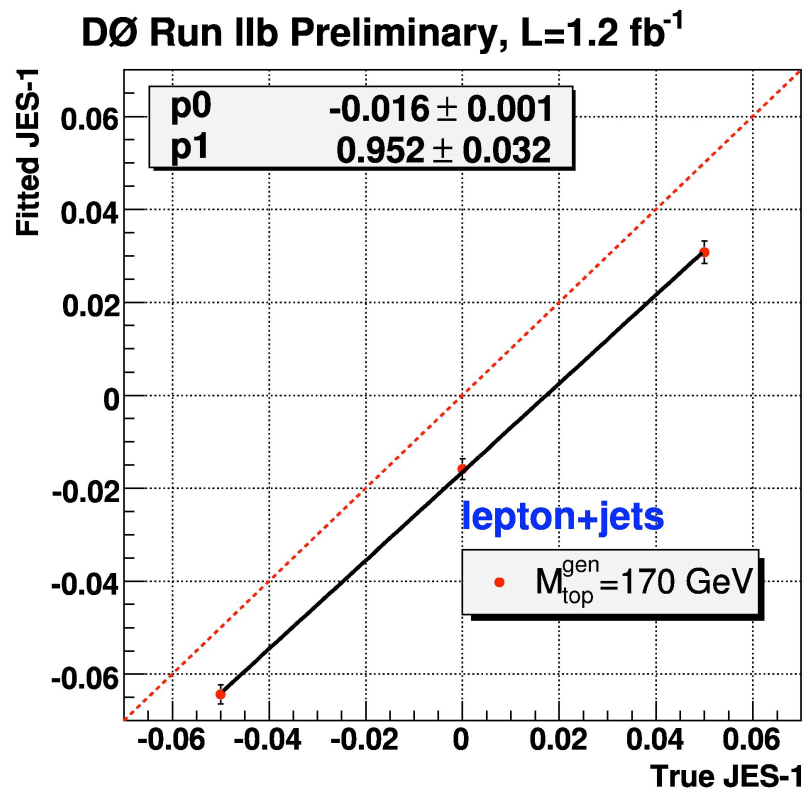
$$P_{sig}(x; m_{top}, JES) = \underbrace{Acc(x)}_{\text{Acceptance (selection, trigger,...)}} \times \frac{1}{\sigma} \int d^n \underbrace{\sigma(y; m_{top})}_{\text{LO-Matrix element x phase space}} \underbrace{dq_1\, dq_2\, f(q_1)\, f(q_2)}_{\text{PDF's}} \underbrace{W(x, y; JES)}_{\text{Transfer Functions (Probability to measure x when y was produced)}}$$

- **maximum Likelihood fit:**

$$L(x_1, .., x_n; m_{\text{top}}, JES, f_{\text{top}}) = \prod_{i=1}^{n} P_{\text{evt}}(x_i; m_{\text{top}}, JES, f_{\text{top}})$$

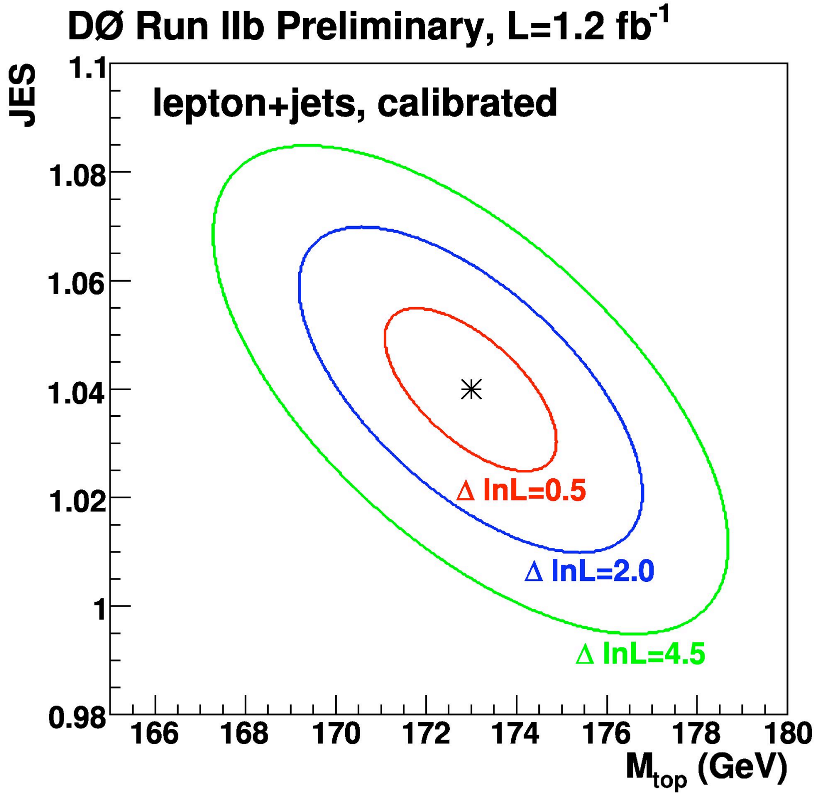
Check you get the right answer

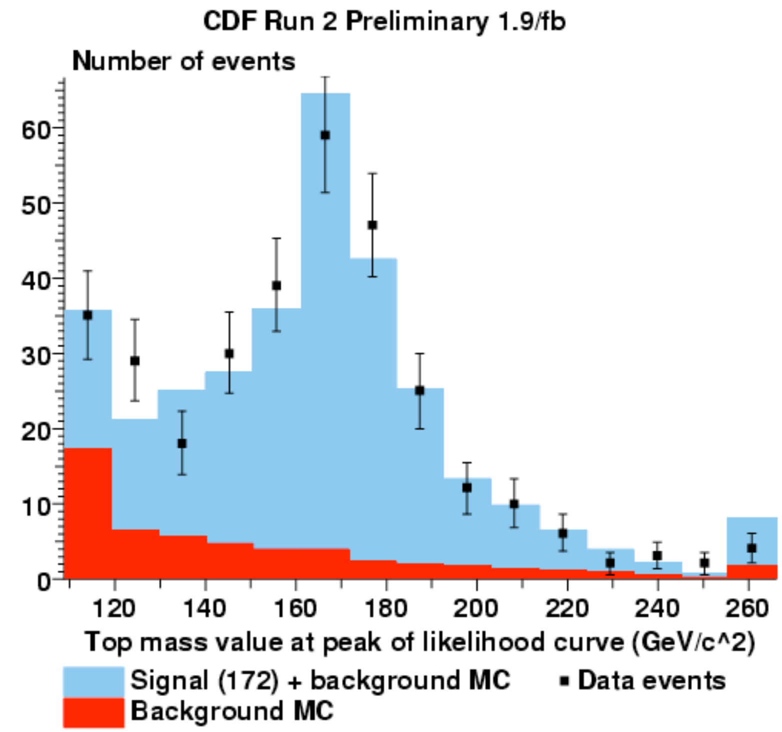




- Run "Pseudo-Experiments" on Monte Carlo to see if you get out the mass that was put in
  - Pretend MC is data and run analysis on it N times
- Non-trivial cross check given the complexity of the method
  - If not: derive "calibration curve" from slope and offset

Matrix Element Top Mass Results





DØ: 2.2 fb^{-1}

m_{top} = 172.2 ± 1.0 (stat) ± 1.4 (syst) GeV

±1.0%

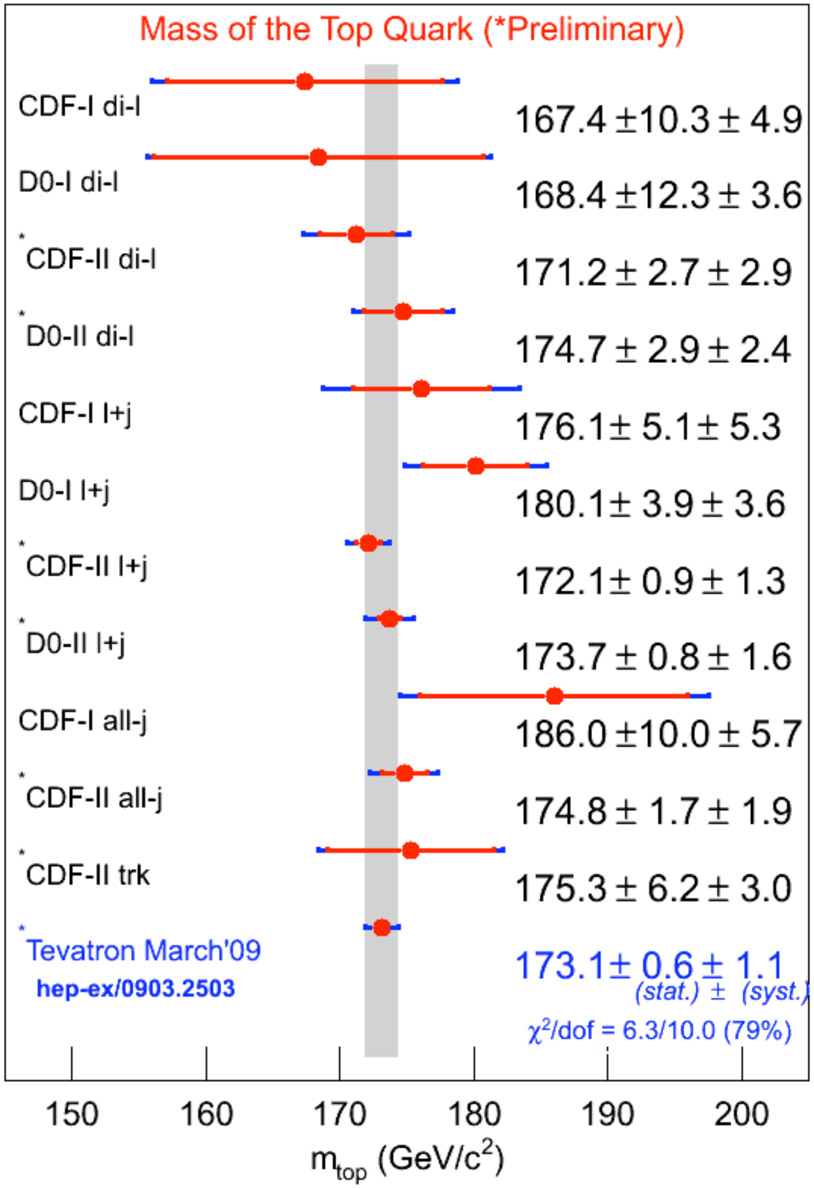
CDF: 2.9 fb^{-1}

m_{top} = 172.2 ± 1.0 (stat) ± 1.3 (syst) GeV

±1.0%

Combining M_{top} Results

- Excellent results in each channel
  - Dilepton
  - Lepton+jets
  - All-hadronic
- Combine them to improve precision
  - Include Run-I results
  - Account for correlations
- Uncertainty: 1.2 GeV
  - Dominated by systematic uncertainties



LHC Perspectives

Systematic uncertainty	1 b-tagged jet	No b-tagging
Light jet energy scale	0.3 GeV/%	0.4 GeV/%
b jet energy scale	0.7 GeV/%	0.7 GeV/%
ISR/FSR	$\simeq 0.4$ GeV	$\simeq 0.4$ GeV
b quark fragmentation	≤ 0.1 GeV	≤ 0.1 GeV
Background	< 1 GeV	1 GeV

- Precisions similar/better to Tevatron when detector understood with ~100 pb^{-1}